

From: Ibrahim, Medina A.
Sent: Monday, September 12, 2005 11:02 AM
To: STIC-Biotech/ChemLib
Subject: 10/713,381

Please search the following:

1. SEQ ID NO: 1 and 2.
2. the sequence from nucleotide position 1 to nucleotide position 1311 of SEQ ID NO: 1 or 2.
3. the sequence from nucleotide position 1155 to nucleotide position 1311 of SEQ ID NO: 1 or 2.
4. the sequence from nucleotide position 1179 to nucleotide position 1208 of SEQ ID NO: 1 or 2.
5. the sequence from nucleotide position 1239 to nucleotide position 1278 of SEQ ID NO: 1 or 2.
6. an oligo of SEQ ID NO: 1 or 2.

Please search both commercial and patent (including pending) databases. Thanks

Medina A. Ibrahim
Patent Examiner, GAU-1638
transgenic plants and plant breeding
Remsen-2B07
mailbox-2C18
(571) 272-0797

RECEIVED
SEP 12 2005
STIC-BIOTECH/CHEM. DIVISION
(STIC)

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2- _____
Date Searcher Picked up: 9/13/05
Date Completed: _____
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA#: 17 AA#: _____
Interference: _____ SPDI: _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure#: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable:

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other(Specify): _____

This page blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 08:25:31 ; Search time 4399.14 Seconds

14440.280 Million cell updates/sec

Title: US-10-713-381-1_COPY_1_1311
Project: 1311

Sequence: 1 ccattggtgctctctatgaaa.....cgcacatattgttctcctgaac 1311

Scoring table: IDENTITY_NUC

Searched: 4708233 beq, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Maximum DB seq length: 2000000000

Post-processing: Minimum Match of 100%

Listing first 45 summaries

Database

```

1:  GenEmbl:*
2:  gb_ha:*
3:  gb_hcg:*
4:  gb_in:*
5:  gb_om:*
6:  gb_ov:*
7:  gb_pat:*
8:  gb_ph:*
9:  gb_pl:*
10: gb_pr:*
11: gb_ro:*
12: gb_stc:*
13: gb_sy:*
14: gb_un:*
15: gb_vl:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1311	100.0	1394	6	AX224394	AX224394 Sequence
2	1311	100.0	1394	6	AX224395	AX224395 Sequence
3	1311	100.0	1394	6	BD062176	BD062176 Male tis
4	1311	100.0	1394	6	BD062177	BD062177 Male tis
5	1311	100.0	3343	8	AF360356	AF360356 Zea mays
6	160.6	12.3	255	6	AX224402	AX224402 Sequence
7	146	11.1	158	6	AX224396	AX224396 Sequence
8	77	5.9	186159	2	AC147602	AC147602 Zea mays
9	73.4	5.6	25769	3	AF034389	AF034389 Dictyostel
10	71.2	5.4	713	3	AF034389	AF034389 Dictyostel
11	71.2	5.4	3576	3	BD065662	BD065662 Dictyostel
12	70.8	5.4	149526	2	BX957346	BX957346 Danto re
13	70.4	5.4	347582	3	PFMAL4P1	AL034557 Plasmodi
14	70.2	5.4	883	6	CQ422641	CQ422641 Sequence
15	70	5.3	241	11	BV119878	BV119878 Plasmodi
16	70	5.3	247	11	BV119874	BV119874 Plasmodi
17	70	5.3	248	11	BV119877	BV119877 Plasmodi
18	69.8	5.3	13400	6	CQ855052	CQ855052 Sequence
19	69.6	5.3	139583	2	AC151050	AC151050 Zea mays

C	20	69.4	5.3	110000	8	CR382132_17	Constitution (18 o
C	21	69.4	5.3	181381	2	CR450818	Danio rer
C	22	69.2	5.3	114817	2	AC149836	AC149836 Zea mays
C	23	69.2	5.3	266544	3	AC116556	AC116556 Dicotyle
C	24	69.2	5.3	164443	2	CR751821	CR751821 Danio rer
C	25	68.8	5.2	7347	1	AF211124	AF211124 Carboneil
C	26	68.6	5.2	186199	2	AC147602	AC147602 Zea mays
C	27	68.4	5.2	213	11	BV119875	BV119875 PZAO1377
C	28	68.4	5.2	216	11	BV119871	BV119871 PZAO1377
C	29	68.4	5.2	224	11	BV119882	BV119882 PZAO1377
C	30	68.4	5.2	245	11	BV119881	BV119881 PZAO1377
C	31	68.4	5.2	246	11	BV119869	BV119869 PZAO1377
C	32	68.2	5.2	346296	8	AF090447	AF090447 Zea mays
C	33	68	5.2	110000	2	PFMAL7P1_02	Constitution (3 of
C	34	67.8	5.2	227495	2	AC148480	AC148480 Zea mays
C	35	67.6	5.2	351	11	BV110530	BV110530 PZAO1059
C	36	67.6	5.2	331039	3	AC116579	AC116579 Dicotyle
C	37	67.4	5.1	170	11	BV119873	BV119873 PZAO1377
C	38	67	5.1	2781	3	AT044085	AT044085 Dicotyle
C	39	67	5.1	110000	3	AC116584_1	Constitution (2 of
C	40	66.8	5.1	228	11	BV119870	BV119870 PZAO1377
C	41	66.8	5.1	234	11	BV119883	BV119883 PZAO1377
C	42	66.8	5.1	7774	12	AF338825	AF338825 Cloning
C	43	66.8	5.1	8658	12	AF338824	AF338824 Cloning
C	44	66.8	5.1	12029	3	AE001434	AE001434 Plasmodu
C	45	66.8	5.1	47573	3	AF030694	AF030694 Plasmodu

ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	AX224394	1394 bp	DNA	linear	PAT 10-SEP-2001
AX224394	Sequence 1 from Patent WO0160997.						
ACCESSION	AX224394						
VERSION	AX224394.1	GI:15554636					
KEYWORDS							
SOURCE							
ORGANISM							
	Zea mays						
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.						
REFERENCE							
AUTHORS	Albertsen,M.C., Fox,T.W., Garnaat,C.W., Hufman,G. and Kendall,T.L.						
TITLE	Male tissue-preferred regulatory region and method of using same						
JOURNAL	Patent: WO 0160997-A 1 23-AUG-2001; HI-BRED INTERNATIONAL, INC. (US)						
FEATURES							
Source	Location/Qualifiers						
	1..1394						
	/organism="Zea mays"						
	/mol_type="unassigned DNA"						
	/db_xref="taxon:4577"						
ORIGIN							
Query Match	100.0%; Score 1311; DB 6; Length 1394;						
Best Local Similarity	100.0%; Freq. No. 3.6e-52;						
Matches 1311; Conservative	0; Mismatches 0; Indels 0; Gaps 0;						
Qy	1 CCATGGTGTCCTATGAAAAAGATGATACATGTCGTATATCCGTTTCTTAGGGTCC 60						
Db	1 CCAATGGTGTCCTATGAAAAAGATGATGACATGTCGTATATCCGTTTCTTAGGGTCC 60						
Qy	61 CTTTCTTTCGCTTATTACTGATGTAATCGGGGTTACAAAACTTCCACGGGTGCATGAT 120						
Db	61 CTTTCTTTCGCTTATTACTGATGTAATCGGGGTTACAAAACTTCCACGGGTGCATGAT 120						
Qy	121 CTCGATGTTCCACTTCCACCTCGGGGTTGCACATTTCTTGATGTCGGGTGTTCCAT 180						
Db	121 CTCGATGTTCCACTTCCACCTCGGGGTTGCACATTTCTTGATGTCGGGTGTTCCAT 180						
Qy	181 CTGACGAGGCCCATCAGACACTTTTGGGACACCCATCAAGGGCTTTTGGATGGCCCA 240						

Db	181	CTGACCGAGGCCCATTCAGACACCTTTGGGACACCCATCAGGGCCTTTGGATGGCCCA	240
Qy	241	CGAGACGTATCGGGTCGTGGTGTATCCAGGGGATATATGTCCCCACATCGTCACTTATA	300
Db	241	CGAGACGTATCGGGTCGTGGTGTATCCAGGGGATATATGTCCCCACATCGTCACTTATA	300
Qy	301	TTATTATCTTTAGATATTATTTAATTTTGGAAAAATACAAACTTATACCTTTGTGTAT	360
Db	301	TTATTATCTTTAGATATTATTTAATTTTGGAAAAATACAAACTTATACCTTTGTGTAT	360
Qy	361	GGGGCCCTCAGCATAGATTTTGGCTTAGGGCCGGAATATGGAGAGACAGCATGTCTAAGTG	420
Db	361	GGGGCCCTCAGCATAGATTTTGGCTTAGGGCCGGAATATGGAGAGACAGCATGTCTAAGTG	420
Qy	421	TCCACTATTTGGGACTACCCAGAACAGATTTAAAAAATATACCAAGTAACATACTACACT	480
Db	421	TCCACTATTTGGGACTACCCAGAACAGATTTAAAAAATATACCAAGTAACATACTACACT	480
Qy	481	CGAAAGCTATCATTGTATGTTTAAAGAAAATCTATATTAACACGATCCTCTTAAAAAA	540
Db	481	CGAAAGCTATCATTGTATGTTTAAAGAAAATCTATATTAACACGATCCTCTTAAAAAA	540
Qy	541	CAAGGATATTTGGAAAGAGACAAATATATGTACAGTTTCAAAACATCTAAGGCGGCAAA	600
Db	541	CAAGGATATTTGGAAAGAGACAAATATATGTACAGTTTCAAAACATCTAAGGCGGCAAA	600
Qy	601	TTATATCGAAAGGTAAAGCTATGACGTTCCAGATTTTCTTTTCATCTCTGTATTTGTT	660
Db	601	TTATATCGAAAGGTAAAGCTATGACGTTCCAGATTTTCTTTTCATCTCTGTATTTGTT	660
Qy	661	ATTGTTTATATATACATTTTCTTCTCTTACATATAGATATTTCTTCCGATTTTATAAA	720
Db	661	ATTGTTTATATATACATTTTCTTCTCTTACATATAGATATTTCTTCCGATTTTATAAA	720
Qy	721	ATGACATATAAGCATTTTATATAAGACAGCCAGTGTGTAGATCTCGTCCAAAAATC	780
Db	721	ATGACATATAAGCATTTTATATAAGACAGCCAGTGTGTAGATCTCGTCCAAAAATC	780
Qy	781	TTTCTGATTTTTTTTAAAGACTAGTTTGGCAACCCGTCTTTCTTCAAGAATTTTGATTT	840
Db	781	TTTCTGATTTTTTTTAAAGACTAGTTTGGCAACCCGTCTTTCTTCAAGAATTTTGATTT	840
Qy	841	TTCAAAAAAAATTTAGTTTATTTTCTCTTATATAAATATAGAAAACTTAGAAAAATAGAT	900
Db	841	TTCAAAAAAAATTTAGTTTATTTTCTCTTATATAAATATAGAAAACTTAGAAAAATAGAT	900
Qy	901	TGCCGACACTAGGCCCTAGATGTTCCTCCCAATAATTACATCATCTGTATATATTATTTG	960
Db	901	TGCCGACACTAGGCCCTAGATGTTCCTCCCAATAATTACATCATCTGTATATATTATTTG	960
Qy	961	GCGACGCCCATTAATTAATTAAACCGAAACTGAAATCGAGCGGAAACCAATCTGACCTAT	1020
Db	961	GCGACGCCCATTAATTAATTAAACCGAAACTGAAATCGAGCGGAAACCAATCTGACCTAT	1020
Qy	1021	TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAATCAGTTTAAATGCAATGTGCC	1080
Db	1021	TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAATCAGTTTAAATGCAATGTGCC	1080
Qy	1081	TGAGATGTGGCGGTTTGGCAACGATATGCCACCGTAAATCAATAGCTCATAGGTGTCTTAGTGA	1140
Db	1081	TGAGATGTGGCGGTTTGGCAACGATATGCCACCGTAAATCAATAGCTCATAGGTGTCTTAGTGA	1140
Qy	1141	GGTTTGGGCGACTCTCGTGTCACTCTCAATGGCACTAATGTGTTTCAACCGTTTCTGTC	1200
Db	1141	GGTTTGGGCGACTCTCGTGTCACTCTCAATGGCACTAATGTGTTTCAACCGTTTCTGTC	1200
Qy	1201	TTGTTTCATGTGCCAAGCCTTGCTATTTCTGAACCAAGAGATACCTACTCCCAACAAT	1260
Db	1201	TTGTTTCATGTGCCAAGCCTTGCTATTTCTGAACCAAGAGATACCTACTCCCAACAAT	1260
Qy	1261	CCATCTTACTATGCAACTTCCATGGAACAGCACTATATGTTTCTGTAAAC	1311
Db	1261	CCATCTTACTATGCAACTTCCATGGAACAGCACTATATGTTTCTGTAAAC	1311

LOCUS	AX224395	1394 bp	DNA	linear	PAT 10-SEP-2001
DEFINITION	Sequence 2 from Patent WO0160997.				
ACCESSION	AX224395				
VERSION	AX224395.1	GI:15554637			
KEYWORDS					
SOURCE					
ORGANISM	Zea mays				
REFERENCE					
AUTHORS	1				
TITLE	Albertsen, M.C., Fox, T.W., Garnat, C.W., Huffman, G. and Kendall, T.L.				
JOURNAL	Male tissue-preferred regulatory region and method of using same				
FEATURES	Patent: WO 0160997-A 2 23-AUG-2001.				
source	PIONEER HI-BRED INTERNATIONAL, INC. (US)				
	Location/Qualifiers				
	1. 1394				
	/organism="Zea mays"				
	/mol_type="unassembled DNA"				
	/db_xref="taxon:4577"				
ORIGIN					
Query Match	100.0%; Score 1311; DB 6; Length 1394;				
Best Local Similarity	100.0%; Pred. No. 3,6e-252;				
Matches 1311; Conservative	0; Mismatches 0; Indels 0; Gaps 0;				
QY	1	CCATGCTGTCTCTATGAAAAAGATGAGTACAAATGTCTATATCCGTTTCTTAAGGTC	60		
DB	1	CCATGCTGTCTCTATGAAAAAGATGAGTACAAATGTCTATATCCGTTTCTTAAGGTC	60		
QY	61	CTTCTTTCGCTTATTACTGATCGAATCGGGGTTACAAAAAATTCCACGGGTGCATGAT	120		
DB	61	CTTCTTTCGCTTATTACTGATCGAATCGGGGTTACAAAAAATTCCACGGGTGCATGAT	120		
QY	121	CTCATGTTCCACTCTTCCACCTCGGCTTGCACTTTCTTGATGTGGGTGTC	180		
DB	121	CTCATGTTCCACTCTTCCACCTCGGCTTGCACTTTCTTGATGTGGGTGTC	180		
QY	181	CTGACGAGGCGCATCAAGACCTTTCGGGACACCCATCAAGGGCGTTTGGATGGCCCA	240		
DB	181	CTGACGAGGCGCATCAAGACCTTTCGGGACACCCATCAAGGGCGTTTGGATGGCCCA	240		
QY	241	CGAGACGATTCGGGTCGTGTGATCCAGGGGATATATGTCCCCACAAATGTCACCTATA	300		
DB	241	CGAGACGATTCGGGTCGTGTGATCCAGGGGATATATGTCCCCACAAATGTCACCTATA	300		
QY	301	TTATATATCTTTAGATATTTATTAATTTTGGAAAAATAAACAATTATCTTTGTGTA	360		
DB	301	TTATATATCTTTAGATATTTATTAATTTTGGAAAAATAAACAATTATCTTTGTGTA	360		
QY	361	GGGCGTCAGCATAGATTTTGGCTTAAAGGCGCCAGAAATGCGAGGACCGCATGTCTAGTG	420		
DB	361	GGGCGTCAGCATAGATTTTGGCTTAAAGGCGCCAGAAATGCGAGGACCGCATGTCTAGTG	420		
QY	421	TCACATATGGCACTACCCAGAACAGATTTTAAAAAATAAACAAGTAATCTAATCCACT	480		
DB	421	TCACATATGGCACTACCCAGAACAGATTTTAAAAAATAAACAAGTAATCTAATCCACT	480		
QY	481	CGAAGGATATCATATATGTTTAAAGAAATCATATTAATAACAACAATCTCTTAAAAAA	540		
DB	481	CGAAGGATATCATATATGTTTAAAGAAATCATATTAATAACAACAATCTCTTAAAAAA	540		
QY	541	CAAGCATATTTTCGAAAGAGCAAAATTAATGTTACAGTTTCAACAAATCTAAGACGACAA	600		
DB	541	CAAGCATATTTTCGAAAGAGCAAAATTAATGTTACAGTTTCAACAAATCTAAGACGACAA	600		
QY	601	TTATATGAAAGGTAGCTATGACGTTACAGATTTTCTTTTCAATCTTGTATATTTGTT	660		
DB	601	TTATATGAAAGGTAGCTATGACGTTACAGATTTTCTTTTCAATCTTGTATATTTGTT	660		

QY 661 ATTGTTTATATACATTTTCTCTTACAAAGAGATTTTCTCCGATTTTATMAA 720
DB 661 ATTGTTTATATACATTTTCTCTTACAAAGAGATTTTCTCCGATTTTATMAA 720
QY 721 ATGACATATAAGTCATTTTATATATAGAGACGCGATCTCGTAGATCTCGTCAAAAATC 780
DB 721 ATGACATATAAGTCATTTTATATATAGAGACGCGATCTCGTAGATCTCGTCAAAAATC 780
QY 781 TTTCTGATTTTTTAAAGCTAGTTGGCAACCCCTGTTCTTCAAAAGAAATTTGATTTT 840
DB 781 TTTCTGATTTTTTAAAGCTAGTTGGCAACCCCTGTTCTTCAAAAGAAATTTGATTTT 840
QY 841 TTCAAAAAAATAGTTATTTTCTCTTATATAAATGAAAAACCTTAGAAAAATAGACT 900
DB 841 TTCAAAAAAATAGTTATTTTCTCTTATATAAATGAAAAACCTTAGAAAAATAGACT 900
QY 901 TGCACAGCTAGCCCTAGAAATGTTTCCCAATTAATCACTGTGTATTAATTTTGG 960
DB 901 TGCACAGCTAGCCCTAGAAATGTTTCCCAATTAATCACTGTGTATTAATTTTGG 960
QY 961 GCCAGCCCAATTAATTTTAAACCGAACTGAATCGAACCAATCTGAGCTAT 1020
DB 961 GCCAGCCCAATTAATTTTAAACCGAACTGAATCGAACCAATCTGAGCTAT 1020
QY 1021 TTCTTAGATTAATTAATAAGGAGAGAGAGAGAAATCACTTTTAATGATTTGCC 1080
DB 1021 TTCTTAGATTAATTAATAAGGAGAGAGAGAGAAATCACTTTTAATGATTTGCC 1080
QY 1081 TGAGATGTGGGTTTGGCAAGATAGCCACCTGATCATAGTGTAGTGGCTACGTCA 1140
DB 1081 TGAGATGTGGGTTTGGCAAGATAGCCACCTGATCATAGTGTAGTGGCTACGTCA 1140
QY 1141 GGTTCGGACGCTCTGTCATCTCAGATGAGCATGATGCTTGTTCACACCTTCTGC 1200
DB 1141 GGTTCGGACGCTCTGTCATCTCAGATGAGCATGATGCTTGTTCACACCTTCTGC 1200
QY 1201 TTGTTCCATGCTGCAAGCCCTTGGCTATTTCTGAAACCAAGAGATTAACCTCCCAACAT 1260
DB 1201 TTGTTCCATGCTGCAAGCCCTTGGCTATTTCTGAAACCAAGAGATTAACCTCCCAACAT 1260
QY 1261 CCATCTTACTATGCAACTTCCATGCAACGCACTATGTTTCTCTGAAC 1311
DB 1261 CCATCTTACTATGCAACTTCCATGCAACGCACTATGTTTCTCTGAAC 1311

RESULT 3
BD062176 1394 bp DNA linear PAT 27-AUG-2002
LOCUS BD062176
DEFINITION Male tissue-preferred regulatory region and method of using same.
ACCESSION BD062176.1 GI:22607781
VERSION JP 2001520523-A/1.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1394)
AUTHORS Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G.A. and Kendall,T.L.
TITLE Male tissue-preferred regulatory region and method of using same
JOURNAL Patient: JP 2001520523-A.1 30-OCT-2001;
COMMENT PIONEER HI BRED INTERNATIONAL INC
PN JP 2001520523-A/1
PD 30-OCT-2001
PF 19-JUN-1998 JP 1999504910
PR 23-JUN-1997 US 08/880499
PI MARC C ALBERTSEN, TIMOTHY W FOX, CARL W GARNAT, GARY A HUFFMAN,
PI TIMMY L KENDALL
PC C12N15/82, C12N15/29, C12N9/24, C12N9/22, C12N9/10, C12N9/00 PC
PC A01H5/00
CC Strandedness: Single;

CC Topology: Linear;
FH Key Location/Qualifiers.
FEATURES
Source 1..1394
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 1311; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 3.6e-252;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCATGCTCTCTATGAAAAAGATGACATATGCTCTATATCCGTTTCTTAAAGGCTCC 60
DB 1 CCATGCTCTCTATGAAAAAGATGACATATGCTCTATATCCGTTTCTTAAAGGCTCC 60
QY 61 CTTCCTCTGCTTATTAAGTGAATCGGGGTTAATAAACTTCCAGGGGTCATGAT 120
DB 61 CTTCCTCTGCTTATTAAGTGAATCGGGGTTAATAAACTTCCAGGGGTCATGAT 120
QY 121 CTCATGTTTCACTTCTCCACCTCGCTTGACATTTCTTGAGTCTGGTGGTCCAT 180
DB 121 CTCATGTTTCACTTCTCCACCTCGCTTGACATTTCTTGAGTCTGGTGGTCCAT 180
QY 181 CTGACCGAGGCCCATGACACCTTTGGGACACCATCAAGGGCTTTGGATGGCCCA 240
DB 181 CTGACCGAGGCCCATGACACCTTTGGGACACCATCAAGGGCTTTGGATGGCCCA 240
QY 241 CGAGAGTATCGGGGTGGTGTATCGAGGGGATATATGTCCTCCCAATCTGACCTAT 300
DB 241 CGAGAGTATCGGGGTGGTGTATCGAGGGGATATATGTCCTCCCAATCTGACCTAT 300
QY 301 TTATATTTCTTATGATTAATTTTGGAAAAATACAACTTATATCTTTGTGTA 360
DB 301 TTATATTTCTTATGATTAATTTTGGAAAAATACAACTTATATCTTTGTGTA 360
QY 361 GGGCTCAGATGATTTTTCCTTAAAGGCCCAAAATGCGAGACACGCTATGTAAGT 420
DB 361 GGGCTCAGATGATTTTTCCTTAAAGGCCCAAAATGCGAGACACGCTATGTAAGT 420
QY 421 TCCACTATTTGGGACCTCCAGAAACAGATTTTAAATAATTAACCAAGTAAATCCACT 480
DB 421 TCCACTATTTGGGACCTCCAGAAACAGATTTTAAATAATTAACCAAGTAAATCCACT 480
QY 481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAACACGATCTTTAAAAAA 540
DB 481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAACACGATCTTTAAAAAA 540
QY 541 CAAGCATATTTGAAAGAGACAAATTAATGTAAGTTTCAAAATCTTAAGGCGCAAA 600
DB 541 CAAGCATATTTGAAAGAGACAAATTAATGTAAGTTTCAAAATCTTAAGGCGCAAA 600
QY 601 TTATATGAAAGGTATGATGCTATGAGTTTCTTTTCTTTCATCTGTTATTTTGT 660
DB 601 TTATATGAAAGGTATGATGCTATGAGTTTCTTTTCTTTCATCTGTTATTTTGT 660
QY 661 ATTGTTTATATACATTTTCTCTTCAATAGAGATTTTCTCCGATTTTAA 720
DB 661 ATTGTTTATATACATTTTCTCTTCAATAGAGATTTTCTCCGATTTTAA 720
QY 721 ATGACATATAAGTCATTTTATATATAGAGACGCGATCTCGTAGATCTCGTCAAAAATC 780
DB 721 ATGACATATAAGTCATTTTATATATAGAGACGCGATCTCGTAGATCTCGTCAAAAATC 780
QY 781 TTTCTGATTTTTTAAAGCTAGTTGGCAACCCCTGTTCTTCAAAAGAAATTTGATTTT 840
DB 781 TTTCTGATTTTTTAAAGCTAGTTGGCAACCCCTGTTCTTCAAAAGAAATTTGATTTT 840
QY 841 TTCAAAAAAATAGTTATTTTCTCTTATATAAATGAAAAACCTTAGAAAAATAGACT 900
DB 841 TTCAAAAAAATAGTTATTTTCTCTTATATAAATGAAAAACCTTAGAAAAATAGACT 900

QY 901 TGCAGACTAGCCCTAGAAATGTTTCCCAATAAATTAACAATCACTGTGTATTAATTG 960
| | | | |
Db 901 TGCAGACTAGCCCTAGAAATGTTTCCCAATAAATTAACAATCACTGTGTATTAATTG 960
QY 961 GCCAGCCCCATAATTATTAAACCGAACTGAATCGAGCAACCAAACTGAGCTAT 1020
| | | | |
Db 961 GCCAGCCCCATAATTATTAAACCGAACTGAATCGAGCAACCAAACTGAGCTAT 1020
QY 1021 TTCTAGATTATTAATAAAGGAGAGAGAGAAATCACTTTTAAGTCAATTCCTCC 1080
| | | | |
Db 1021 TTCTAGATTATTAATAAAGGAGAGAGAGAAATCACTTTTAAGTCAATTCCTCC 1080
QY 1081 TGAGATGTGCGGTTTGCGAAGATAGCAACCGTATCATAGTGTGCTTACGTCA 1140
| | | | |
Db 1081 TGAGATGTGCGGTTTGCGAAGATAGCAACCGTATCATAGTGTGCTTACGTCA 1140
QY 1141 GGTTCGGAGCTCTCGTGTATCTCAGATGGCATPACTACATGCTTGTCAACCGTTCGTC 1200
| | | | |
Db 1141 GGTTCGGAGCTCTCGTGTATCTCAGATGGCATPACTACATGCTTGTCAACCGTTCGTC 1200
QY 1201 TTGTTCCATGTCGCAAGCCTTGCTATTCTGAACCAAGAGATACCTACCTCCAAACAT 1260
| | | | |
Db 1201 TTGTTCCATGTCGCAAGCCTTGCTATTCTGAACCAAGAGATACCTACCTCCAAACAT 1260
QY 1261 CCATCTTACTCATGCAACTTCATGCAACACGCGACATATGTTTCTGTAAC 1311
| | | | |
Db 1261 CCATCTTACTCATGCAACTTCATGCAACACGCGACATATGTTTCTGTAAC 1311
RESULT 4
BD062177 1394 bp DNA linear PAT 27-AUG-2002
LOCUS BD062177 Male tissue-preferred regulatory region and method of using same.
DEFINITION BD062177
ACCESSION BD062177.1 GI:22607782
VERSION JP 2001520523-A/2.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 1394)
Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G.A. and
Kendall,I.T.L.
Male tissue-preferred regulatory region and method of using same
Patient: JP 2001520523-A 2 30-OCT-2001;
PIONEER HI BRED INTERNATIONAL INC
COMMENT EN JP 2001520523-A/2
PD 30-OCT-2001 JP 1999504910
PF 19-JUN-1998 JP 1999504910
PI 23-JUN-1997 US 08/880499
PR MARC C ALBERTSEN,TIMOTHY W FOX,CARL W GARNAAAT,GARY A HUFFMAN,
PI TIMMY L KENDALL
PC C12N15/82,C12N15/29,C12N9/24,C12N9/22,C12N9/10,C12N9/00 PC
,C07K14/34,C1201/68,
PC A01H5/00
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers.
FEATURES
source 1.1394
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 1311; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 3, 6e-252;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 CCATGCTCTCATGATAAAGATGAGTCAATGTGTATTCCTTTCTTAAAGGTCC 60
1 CCAATGCTCTCATGATAAAGATGAGTCAATGTGTATTCCTTTCTTAAAGGTCC 60

QY 61 CTCTCTCTGCTTATTAAGTGAATCGGGTTACAAAAAACTTCCAGCGGTGATGAT 120
| | | | |
Db 61 CTCTCTCTGCTTATTAAGTGAATCGGGTTACAAAAAACTTCCAGCGGTGATGAT 120
QY 121 CTCATGTTTCACTTCTCCACCTCGGCTTGCACTTTCTTGATGTGGTGGTCCAT 180
| | | | |
Db 121 CTCATGTTTCACTTCTCCACCTCGGCTTGCACTTTCTTGATGTGGTGGTCCAT 180
QY 181 CTGACCGAGCCCATCAAGACACTTTCCGGGAACCCATCAAGGGCTTTCCGATGGCCCA 240
| | | | |
Db 181 CTGACCGAGCCCATCAAGACACTTTCCGGGAACCCATCAAGGGCTTTCCGATGGCCCA 240
QY 241 CGAGACGTATCGGGTGTGATGATCAGGGAGTATATGTCCCCCAATCGTCACTATA 300
| | | | |
Db 241 CGAGACGTATCGGGTGTGATGATCAGGGAGTATATGTCCCCCAATCGTCACTATA 300
QY 301 TTATTAATCTTTAGATATTAATTTTGGAAAAATTAACAACCTTATATCTTTGTGTA 360
| | | | |
Db 301 TTATTAATCTTTAGATATTAATTTTGGAAAAATTAACAACCTTATATCTTTGTGTA 360
QY 361 GGGGCTCAGCATAGATTTTGGTTAGGGCCGGAATGCGAGACAGCATGTCTAGTG 420
| | | | |
Db 361 GGGGCTCAGCATAGATTTTGGTTAGGGCCGGAATGCGAGACAGCATGTCTAGTG 420
QY 421 TCCACTATTGGGCACTACCGAGAACAGATTTAAAAAAATTAACCAAGTAACTAATCACT 480
| | | | |
Db 421 TCCACTATTGGGCACTACCGAGAACAGATTTAAAAAAATTAACCAAGTAACTAATCACT 480
QY 481 CGAAAGCTATCATGTATGTTTAAAGAAACATCTATTAACCAAGATCTCTTAAAAA 540
| | | | |
Db 481 CGAAAGCTATCATGTATGTTTAAAGAAACATCTATTAACCAAGATCTCTTAAAAA 540
QY 541 CAAGCATATTTGGAAGAGACAAATTAATGTTACAGTTTACAAACATCTAAGCGGACAA 600
| | | | |
Db 541 CAAGCATATTTGGAAGAGACAAATTAATGTTACAGTTTACAAACATCTAAGCGGACAA 600
QY 601 TTATATCGAAAGATAGCTATGATGATGATGATGATGATGATGATGATGATGATGAT 660
| | | | |
Db 601 TTATATCGAAAGATAGCTATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 661 ATTTGTTTATATACATTTTCTTCTTCAATAGATGATGATGATGATGATGATGATGAT 720
| | | | |
Db 661 ATTTGTTTATATACATTTTCTTCTTCAATAGATGATGATGATGATGATGATGATGAT 720
QY 721 ATGACATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
| | | | |
Db 721 ATGACATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 781 TTCTGATTTTATTAAGAGCTAGTTTGGCAACCTGTCTTCTTCAAGATTTTGAATTT 840
| | | | |
Db 781 TTCTGATTTTATTAAGAGCTAGTTTGGCAACCTGTCTTCTTCAAGATTTTGAATTT 840
QY 841 TTCAAAAAAATTAAGTTATTTTCTTCTTATTAATTAAGAAACCTTAAAGATTAAGAT 900
| | | | |
Db 841 TTCAAAAAAATTAAGTTATTTTCTTCTTATTAATTAAGAAACCTTAAAGATTAAGAT 900
QY 901 TGCAGACTAGCCCTAGAAATGTTTCCCAATAAATTAACAATCACTGTGTATTAATTG 960
| | | | |
Db 901 TGCAGACTAGCCCTAGAAATGTTTCCCAATAAATTAACAATCACTGTGTATTAATTG 960
QY 961 GCCAGCCCCATAATTATTAAACCGAACTGAATCGAGCAACCAAACTGAGCTAT 1020
| | | | |
Db 961 GCCAGCCCCATAATTATTAAACCGAACTGAATCGAGCAACCAAACTGAGCTAT 1020
QY 1021 TTCTAGATTATTAATAAAGGAGAGAGAGAAATCACTTTTAAGTCAATTCCTCC 1080
| | | | |
Db 1021 TTCTAGATTATTAATAAAGGAGAGAGAGAAATCACTTTTAAGTCAATTCCTCC 1080
QY 1081 TGAGATGTGCGGTTTGCGAAGATAGCAACCGTATCATAGTGTGCTTACGTCA 1140
| | | | |
Db 1081 TGAGATGTGCGGTTTGCGAAGATAGCAACCGTATCATAGTGTGCTTACGTCA 1140
QY 1141 GGTTCGGAGCTCTCGTGTATCTCAGATGGCATPACTACATGCTTGTCAACCGTTCGTC 1200
| | | | |

Db	1141	GGTTGGGACCTCTCGTGTATCTACATGACATCTACTGCTTGTTCACACCGTTGCT	1200
Qy	1201	TTGTTCCATGCTCCAAACCTTGGCTATCTGAACCAAGAGATACCTACTCCCAACAT	1260
Db	1201	TTGTTCCATGCTCCAAACCTTGGCTATCTGAACCAAGAGATACCTACTCCCAACAT	1260
Qy	1261	CCATCTTACTACGCAACCTTCCATGCAACACGACATATGTTTCCGAAC	1311
Db	1261	CCATCTTACTACGCAACCTTCCATGCAACACGACATATGTTTCCGAAC	1311
RESULT 5			
LOCUS	AF360356	3343 bp	DNA
DEFINITION	Zea mays male fertility protein (M45) gene, complete cds.		linear PLN 12-MAY-2001
ACCESSION	AF360356		
VERSION	AF360356.1	GI:14028756	
KEYWORDS			
SOURCE	Zea mays		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.		
REFERENCE	1 (bases 1 to 3343)		
AUTHORS	Fox,T.W., Trimmell,M.R. and Albertsen,M.C.		
TITLE	Cloning of M45, a gene required for male fertility from Zea mays		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 3343)		
AUTHORS	Fox,T.W., Trimmell,M.R. and Albertsen,M.C.		
TITLE	Direct Submission		
JOURNAL	Submitted (13-MAR-2001) Trait and Technology Development, Pioneer Hi-Bred Int., Inc., 7300 N.W. 62nd Ave., P.O. Box 1004, Johnston, IA 50131-1004, USA		
FEATURES			
source	location/Qualifiers		
	1..3343		
	/organism="Zea mays"		
	/mol_type="genomic DNA"		
	/cultivar="B73"		
	/db_xref="taxon:4577"		
	/chromosome="9L"		
gene	<1392..32942		
	/gene="M45"		
mRNA	join(<1392..1768,1898..2182,2280..2447,2534..>2942)		
	/gene="M45"		
	/product="male fertility protein"		
	join(1392..1768,1898..2182,2280..2447,2534..2942)		
	/gene="M45"		
	/codon_start=1		
	/product="male fertility protein"		
	/protein_id="PAK52489.1"		
	/db_xref="GI:14028757"		
	/translation="MEKRLQMRGRDGIQVPHLPFAALALALVADPFGSLPAEVDYRIVKHELAPEGVMSNPRDASRLRRLRLEFGEVGEVSPESIEFDLGRGVAGLADGRVAVMGEAEAMETPAYNPNPMSEEVAGNSTRQHEERESCGRLRFGETGLVYADATYGCATVVGSGGCAVSVAAREADDPPIRPANDLDVHRGVSFFIDTSRYSRKDLNILLBEGSGTGLRLRYDPTSGVHVLLKGLVFPNGVQISDHOFLRSETTSKRLRMVLEGRAGEVEFANLPGFPNVNSNGRGQFVAIDCCTPAQOEVFARPMRLTLYFKPLSLKVLTKAKARMTVTLALLDOEGRVVEVLEDRGHEVKKLVSEVVERKLGWTVAAHHNIATIPYLED"		
ORIGIN			
Query Match	100.0%;	Score 1311;	DB 8; Length 3343;
Best Local Similarity	100.0%;	Pred. No. 3.5e-252;	
Matches 1311;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
Qy	1	CCATGCGCTCTATGAAAAAGATGATGATGCTATATCCGTTTCTTAAAGGTC	60
Db	1	CCATGCGCTCTATGAAAAAGATGATGATGCTATATCCGTTTCTTAAAGGTC	60
Qy	61	CTTCTTCGCTTATTAAGTGAATCGGGTTTACAAAACTTCAACGGGTGCATGAT	120
Db	61	CTTCTTCGCTTATTAAGTGAATCGGGTTTACAAAACTTCAACGGGTGCATGAT	120

QY 1201 TTGTTCCATGCTCAAGCCTTGCTATTCTGAACAAGAGATACCTATCCCAACAAT 1260
|||||
Db 1201 TTGTTCCATGCTCAAGCCTTGCTATTCTGAACAAGAGATACCTATCCCAACAAT 1260
|||||
QY 1261 CCATCTTACTCATGCAACTTCATGCAACACGACATATGTTTCTGAAC 1311
|||||
Db 1261 CCATCTTACTCATGCAACTTCATGCAACACGACATATGTTTCTGAAC 1311
|||||
RESULT 6
AX224402 AX224402 255 bp DNA linear PAT 10-SEP-2001
LOCUS Sequence 9 from Patent WO0160997.
ACCESSION AX224402
VERSION AX224402.1 GI:15554644
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
AUTHORS Albertsen, M.C., Fox, T.W., Garnaat, C.W., Huffman, G. and Kendall, T.L.
TITLE Male tissue-preferred regulatory region and method of using same
JOURNAL Patent: WO 0160997-A 9 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
source 1.255
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"
ORIGIN
Query Match 12.3%; Score 160.6; DB 6; Length 255;
Best Local Similarity 97.6%; Pred. No. 5.9e-22;
Matches 163; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1145 CGGCACTCTGCTGTCATCTCACATGGCATACATAGTGTTCACCGTTGCTGTG 1204
|||||
Db 5 CGCGGATCCCGGTATCTACATGGCATACATAGTGTTCACCGTTGCTGTG 64
|||||
QY 1205 TCCATGCTCCAGCCTTGCTTATTCGAACAAGAGATACCTATCCCAACAATCCAT 1264
|||||
Db 65 TCCATGCTCCAGCCTTGCTTATTCGAACAAGAGATACCTATCCCAACAATCCAT 124
|||||
QY 1265 CTTACTCATGCAACTTCATGCAACACGACATATGTTTCTGAAC 1311
|||||
Db 125 CTTACTCATGCAACTTCATGCAACACGACATATGTTTCTGAAC 171
|||||
RESULT 7
AX224396 AX224396 158 bp DNA linear PAT 10-SEP-2001
LOCUS Sequence 3 from Patent WO0160997.
ACCESSION AX224396
VERSION AX224396.1 GI:15554638
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
AUTHORS Albertsen, M.C., Fox, T.W., Garnaat, C.W., Huffman, G. and Kendall, T.L.
TITLE Male tissue-preferred regulatory region and method of using same
JOURNAL Patent: WO 0160997-A 3 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
source 1.158
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"
ORIGIN

Query Match 11.1%; Score 146; DB 6; Length 158;
Best Local Similarity 99.4%; Pred. No. 5e-19;
Matches 157; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1155 CGTTCATCTCATGCAATGCGATCTACATGCTTTCAACCGTTCGTC-TTGTTCCATGCTC 1213
|||||
Db 1 CGTTCATCTCATGCAATGCGATCTACATGCTTTCAACCGTTCGTTGTTCCATGCTC 60
|||||
QY 1214 CAAGCTTGCTTATCTGGAACCAAGAGATACCTATCCCAACAATCTTACTCAT 1273
|||||
Db 61 CAAGCTTGCTTATCTGGAACCAAGAGATACCTATCCCAACAATCTTACTCAT 120
|||||
QY 1274 GCACTTCATGCAACACGACATATGTTTCTGAAC 1311
|||||
Db 121 GCACTTCATGCAACACGACATATGTTTCTGAAC 158
|||||
RESULT 8
AC147602 AC147602 186199 bp DNA linear HTG 17-AUG-2004
LOCUS Zea mays clone ZMMBB0334A01, *** SEQUENCING IN PROGRESS ***, 6
DEFINITION Ordered pieces.
AC147602
AC147602.5 GI:51315585
VERSION HTG; HTGS_PHASE2; HTGS_FULLTOP; HTGS_ACTIVEFIN.
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
AUTHORS Birren, B., Nuebaum, C., Lander, E., Butler, E., Wang, R., Bhatti, A.K.
and Messing, J.
Ze mays, clone ZMMBB0334A01
Unpublished
2 (bases 1 to 186199)
TITLE
JOURNAL
AUTHORS
REFERENCE
AUTHORS Birren, B., Nuebaum, C., Lander, E., Abouelleil, A., Allen, N.,
Anderson, M., Arabchi, H.M., Barna, N., Bastien, V., Bloom, T.,
Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y.,
Collamore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K.,
Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S.,
Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamit, A., Karatae, A., Kells, C., Landers, T., Levine, R.,
Lindblad-Ton, K., Liu, X., Lui, A., Mabbitt, R., Maclean, C.,
Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
Meldrim, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J.,
Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, J., Peterson, K., Phunhbang, P., Pierre, N.,
Rachupka, A., Ramsamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, O., Schauer, S., Schupack, R., Seaman, S., Severy, P., Smith, C.,
Spencer, B., Strange-Thomann, N., Stojanovic, N., Stubbs, M.,
Talamas, J., Testaye, S., Theodore, J., Topham, K., Triviera, M.,
Vasiliiev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zemdek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (19-DEC-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 186199)
TITLE
JOURNAL
AUTHORS
REFERENCE
AUTHORS Arabchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
Boukhalter, B., Camarata, J., Chang, J., Choepel, Y., Collamore, A.,
Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J.S., Dodge, S.,
Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P.,
Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L.,
Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J.,
Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamit, A.,
Karatae, A., Kells, C., Landers, T., Levine, R., Lindblad-Ton, K.,
Liu, X., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P.,
Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J.,


```

TITLE
JOURNAL
COMMENT

Meynen, L., Mhova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,
Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, J., Peterson, K., Phunthong, P., Pierre, N.,
Rachupka, A., Ramasamy, U., Raymond, C., Reta, R., Rise, C., Rogov, P.,
Roman, J., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C.,
Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M.,
Vasilev, H., Venkataraman, V.S., Viet, R., Vo, A., Wilson, B., Wu, X.,
Wynan, D., Young, G., Zainoun, J., Zemdek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (17-AUG-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 17, 2004 this sequence version replaced gi:19658659.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/MIT Center for Genome Research
Center code: WITB
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@broad.mit.edu

-----
Bharit, AK and Messing, J: The Plant Genome Initiative at
Rutgers, Waksman Institute, Rutgers, The State University of New
Jersey, 190 Frelinghuysen Road, Piscataway, NJ 08854, USA
(http://pgir.rutgers.edu)
Butler, E and Wing, R: Arizona Genomics Institute, Biological
Sciences West, 448A, P.O. Box 210088, University of Arizona,
Tucson, AZ 85721, USA (http://www.genome.arizona.edu)

----- Project Information
Center project name: L30003
Center clone name: 334_A1

----- Consensus Information
This consensus is derived from a shotgun assembly that has been
manually curated. It is the best representation of the BAC that we
can generate without further laboratory work. The draft assembly
has been edited, and if possible, ends identified by vector as well
as by BAC end sequences, and contigs ordered and oriented. Bases
that are not Ns are either above Q20 or manually edited. This
assembly was performed with Arachne (Genome Res. 2002 12: 177-189;
Genome Res. 2003 13: 91-96). All trace files for this project are
available at the NCBI trace repository
(http://www.ncbi.nlm.nih.gov/Traces/trace.fcgi?). An exact list of
reads used in this assembly are available at
http://www.broad.mit.edu/annotation/plants/maize/randomclones.html.
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 100617: contig of 100617 bp in length
* 100618 100717: gap of unknown length
* 100718 104730: contig of 4013 bp in length
* 104731 104830: gap of unknown length
* 104831 115104: contig of 10274 bp in length
* 115105 115204: gap of unknown length
* 115205 156396: contig of 41192 bp in length
* 156397 156496: gap of unknown length
* 156497 179936: contig of 23440 bp in length
* 179937 180036: gap of unknown length
* 180037 186199: contig of 6163 bp in length.

Location/Qualifiers
1. 186199
/organism="Zea mays"
/mol_type="Genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ZmBIBC0334A01"
FEATURES
SOURCE

```

ORIGIN	/clone_lib="CHORI-201 Maize B73 BAC Mbl Library"				
Query Match	5.9%	Score 77	DB 2	Length 186199	
Best Local Similarity	71.4%	Pred. No. 2.7e-05			
Matches 115	Conservative 0	Mismatches 45	Indels 1	Gaps 1	
QY	780	CTTTGATGATTTTAAAGAGCTAGTTGGCAACCTGTTCTTTCAAGAAATTTGATTT	839		
DB	162058	CTTTCCTTTTACTTTTACGGCTAGTTGGGAACCTTATTT-TTCTAGAAATTTTCATTT	162116		
QY	840	TTTCAAAAAAATAGTATTATTTCTCTTTTAAATGAAACACTTAGAAAAATPAGG	899		
DB	162117	TTCTTAGGAAATTAATGTTTCATTTCTTTACAAAATGAATTTATTTGTTAAATAGAG	162176		
QY	900	TTGCCAGACTAGCCCTAGATGTTTTCCTCAATTAATTTACAA	940		
DB	162177	TTTCCAAACTACCTCAAGTATTAATGAATGACGAGCA	162217		
RESULT 9					
AC117267		25769 bp	DNA	linear	INV 21-FEB-2004
LOCUS	Dictyostelium discoideum chromosome 2 map 5836255-5862024	strain			
DEFINITION	AX4, complete sequence.				
ACCESSION	AC117267	AC115597			
VERSION	AC117267.2	GI:42733680			
KEYWORDS	HTG.				
SOURCE	Dictyostelium discoideum				
ORGANISM	Dictyostelium discoideum				
REFERENCE	Eukaryota; Mycetozoa; Dictyostelid; Dictyostelium.				
AUTHORS	1 (bases 1 to 25769)				
	Gloeckner,G., Eichinger,J., Szafarski,K., Pachebat,J., Dear,P.,				
	Lehmann,R., Baumgart,C., Pariz,G., April,J.F., Gutigo,R., Kumpf,K.,				
	Tunggal,B., Cox,E., Quail,M.A., Platter,M., Rosenhal,A. and				
	Noegel,A.A.				
	Sequence and analysis of chromosome 2 of Dictyostelium discoideum				
	Nature 418 (6893), 79-85 (2002)				
	12097910				
TITLE	The Dictyostelium Genome Sequencing Consortium				
JOURNAL	2 (bases 1 to 25769)				
AUTHORS	Baumgart,C.				
REFERENCE	Direct Submission				
TITLE	Submitted (09-APR-2002) Genome Analysis, Institute of Molecular				
JOURNAL	Biotechnology, Beutenberg. 11, Jena 07745, Germany				
	On or before Feb 21, 2004 this sequence version replaced				
	gi:119570016, gi:20087114.				
COMMENT	CDS predictions from Genaid may contain errors. Further information				
	is available from IMB Jena, Department of Genome Analysis				
	(http://genome.imb-jena.de/dictyostelium/)				
	and the University Cologne, Institute for Biochemistry I				
	(http://www.uni-koeln.de/dictyostelium/project.shcml				
	Pending				
FEATURES	Agency : Deutsche Forschungsgemeinschaft (DFG).				
source	Location/Qualifiers				
	1..25769				
	/organism="Dictyostelium discoideum"				
	/mol_type="genomic DNA"				
	/strain="AX4"				
	/db_xref="taxon:44689"				
	/chromosome="2"				
	/map="5836255-5862024"				
	complement(join(832..1070,1237..1349,1533..1798,				
	1896..2053,2231..2360))				
	/note="Genaid exon scores (in order of location ranges):				
	23.68, 15.54, 15.75, 9.21, -4.75 - GSCU_ID dd_03587"				
	/codon_start=1				

```
/product="hypothetical protein"
/protein_id="AAS38625.1"
/db_xref="GI:42733681"
/translation="MRLOITLPLSLILRLQKSTRESSNDISTLITSTENFKDYNQ
ILSKSEWFSLEYSSNNKRVFEVDFGNHYEINDRLIKKEFENKRVYVDVSPFCL
NSKFNLEYIKKFTSISRNKKVPMNIESTHENCYNNVLYNNKNNVEOEVLV
FSKQVLSFLQNGYNIIGIKSKNTFKIENNNDKLFSSLPFFYLKNIENOFEBEY
SPLSGMTPDKIKQFKKNDIETILSDVEVNLCSNNNNNNNNNNVDETLTLGSGSG
SGSGSGSGSGAQLSGTGVIINHO"
join(2957..3173,3272..5583)
/note="Gened exon scores (in order of location ranges) :
24.88, 254.96 - GSCJ_ID dd_00753"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAS38626.1"
/db_xref="GI:42733682"
/translation="MEPTKRYVITSNLDKPKEMWTFQSEMSLTDTDTQKSPVNSG
SSSSKGFGEAVLIDLDTAGOEYSAMDQVYRTDCENIVFSIDRSRPFESVQK
OHLERVYDRDVPILIVGNKVDLESRRVSRLEADQLARLVPYIETSAKTSNIEE
AFETLVKTPRNTVYKVVVGGVGKSAIIQIOMHFEVEYDPTIEDSYKQVVIS
GLPPIGSLKKSSSSSSSSSSSGKTGLFNKIFSGDKRQSPQQAASSTIDRGOI
STRRLANVLSYMSNLKSEVPILITDCYCCQCNVILSRPSLVITGPDSTWKEF
CKTSNSNILEGEIIPNKDSVEYVLSPTSTTDSKKEESIYICIVDSGWTGITT
EVPISQSEWNAKGVGASGSPYSIRLECVQSIPMIIDRLIQPKRVVLVTPS
DEVMYITQSNVDGPIVIAGDKLEDPDLIEGRMTYDKLPTASGSDPLAKIKSL
EPVOSTALGPALVLSAIAISOKMSEVICTDGPVNGALIEDLPGAOEFYEKVT
KLAONKTTINIIIGISGSHIDIGVIGKGEOTNGNITIIHLEAREIKRTQNPMTA
TVEMETICAPTELEINKYDSKQSLSRVYKQFPMNLTDLTLISGRNPTFVOQYR
FOQIQTRTKTVRCIRVVAQLOATPDPFTSNANISILAAFTQQAQAKIAQOEY
MESRLHLKATKILRSLCNTDEQEEFYFVEVREMEAPLITICINKQOORVEKATD
DEIOVYKMKXKNSFVEGGRKKDISRRKGEAINKQYNIKFT"
complement(5711..6796)
/note="Gened exon scores (in order of location ranges) :
58.61 - GSCJ_ID dd_00752"
/codon_start=1
/product="similar to Plasmodium falciparum (isolate 3D7).
Hypothetical protein"
/protein_id="AAS38627.1"
/db_xref="GI:42733683"
/translation="MSRYSKSLIERYNGILITERSPRGLINTNNNNNNNN
GTTTTTSGNNNNNSTSPSTNPEVSDRVDSNNQIQIVDSFNINNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
KFNNNNNYFDEFSTNEMITTLNSCTLEQIGTIGNDYSTNTTSPSSSSSTSLAN
KSNHNNMPSLKIINDNNNNISNINEFQIQSNSSSSCSTQCPIDEHIOE
QSLIOEKOLKOIDDDDDDDDDSSPSPPTTTTKTFCLSRPRIKKEEID
QMSRPPKLDKRRPSRDLIPTIANR"
join(8924..9076,9236..10573)
/note="Gened exon scores (in order of location ranges) :
10.37, 128.60 - GSCJ_ID dd_00729"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAS38628.1"
/db_xref="GI:42733684"
/translation="MSSSTITPIPIATTTNNNSNGCHNSKTRTLFKNIKNIENS
PNNAPIKNSVGIIGPNVNNNNKDKNSNDGVNFGKNILPTSPKNSMSTTKHS
SSNSPLAKNEYDIBESIKKELKMDISTGDISPLRESSPISTIDESIATATTK
LTPAFAAASSSSSVKTIIDVLSFKSANTKRPLOIDIVENHOKELIVPLSPTPYN
NNNNNNNNNSOCHNTANGOKYPIFPSPOISPKIAYAOSSSTKTNINNNNNNTPT
TKKANSITTPGOSNNNNHNSNTKSKQNPISPLSLTATPNOKTTTTSKSKSTT
TTPKSNKIKINDVAAFAAVANSTTTTQPAVLIIDGATPSTNESPKTPTLKTEDPKLA
ARQROIDIGKMTAGYNYIALVPKSRKREPTPTPKNQVCSKRSMDQIKKWRQLH
ENYDWTGIFPNEKEVEASNNSTESTLEKMKKLEISENEHNN"
complement(10972..11979)
/note="Gened exon scores (in order of location ranges) :
23.29 - GSCJ_ID dd_00728"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAS38629.1"
/db_xref="GI:42733685"
/translation="MEBKIKFVEIIDPONOKKIOEFTSKLIGKESFTTPKPIVY
EKLWNTQSIIESILKTNGTINKSIFEEIFSYCGYIGERLEBEPPIVFLKPTTGGIL
DDTIFKINSLNKKEKULINSLIYNNKNNKIKGFEPWEIENFQPTIKESFOR
IINANSLWKSSIDSRNINSNSPINSRCSFNEYFKERSDASGDFILITISIGIMDYI
```

```
ENSIKSEFOTINYHAKSFLLINDLYSPNREINENDLNTYKILATOLANSIQSLD
KTIELIDHYLKLSSIEITLKYQNDSTYQLKQVQNSKILSIGIFPAKSKRY
N"
join(12884..13020,13172..15320)
/note="Gened exon scores (in order of location ranges) :
2.42, 136.56 - GSCJ_ID dd_00727"
/codon_start=1
/product="similar to Dictyostelium discoideum (slime
mold). Hypothetical 97.7 kDa protein"
/protein_id="AAS38630.1"
/db_xref="GI:42733686"
/translation="MDNLIDYFLIKREYNQGSFLDCLILSLNGYISYDGLDCMN
FLIENKTVNKEAISYGLANYOQLDEFTLTKKTKCTCKKLIINTHNSLKKCC
GNKKRYSILIIENNNNIINKDYQFNNNKNNNNNNNNNNNNNNNNNNNNNNNN
TLKNERINGDGEITTTTTTTTTSTSTMTTITNGNDNRNLTLPQVARNVVIK
TOILPHALNINHANAKVFLRPIQVQVYKRGYLOSMTLNDHDDDDADESYSDI
DDDDYSOQGNCSSSSSSDIGSSSSINNGISSSSSSILNSLSLBSLSTRVNSL
PEGLKSIFFEKENYVIOQLRLLPSSISIRFSGFNQRLAKGVISDNVISTGDEVTQ
SLDGMNPKQLKHLQFGHKFOQTIKMGQLPSLSLILIDPRYSKVIEIGSIPDESFT
LDYKNSLNSOESISFNPISKSITLVDSFENQNIKANDISNNVLSIHFGEFNS
DIGIKSLNSIREIKRGAFDRDILKCSITSIDQFKNRPILSMOTQTLSTIDGS
KENQIIPGIFHTLKSLNRYHNOIIPADTLPTELSNLGCTNMTQITKDNBYD
CYGISNKGFGSNNSSNFCVGTNGLMEMLNTTSLKTLTNTYNRKIEVDLPSI
ESLNGYHFNQPIGNVLPKLLKPLINSENNQNIADGCIPIGLQTIYIRNSMNF
INSLNPLFTKXINIIDLSHL"
join(16184..16199,17272..17613,17707..19331)
/note="Gened exon scores (in order of location ranges) :
0.08, 32.32, 227.42 - GSCJ_ID dd_00726"
/codon_start=1
/product="similar to Dictyostelium discoideum (slime
mold). Spore coat protein Sp87 precursor (Plj protein)"
/protein_id="AAS38631.1"
/db_xref="GI:42733687"
/translation="MRIFGDCNKTINERECSSRCIYINTPCCGEOKMASCSTPD
TCTYNSCTRNSSNNQVYEVNSKCNELNGLFITTPTEYSGSDAKIKELAKGCSF
QKNSGNSNPTSCCPGEISICEGLSSGSSGTSAGTSAGTSAGTSAGTSAGTS
SSGSSGVSISCTHCEPGEYCSMNVAICLSTGTGGLPTGTSSTAGVSGCLTT
LCPIGHICVEDSNGVCVNGGSGTSGTSGCHDPCGDVCCPGOFHCEGDOG
TAKCVSPPTSSSTSGHPPCKDVTGPDGHCCKQKTKACVSPPTSSSTSGG
NTNPSNNSNCPGPFCECKDKTKACVSGSPQPPKPYCSLRCPNHECRNDQHO
CCVQVHIDRCLRCPHGECKVDHNGKCCVRSRHPPEVCSLRCPHGECKPDHG
KKCCVKIHCDEVDLDCGRPECKIRHGSCKCVASRHPHGHCKNKRCPHGECK
VDQHEKCECVAAHRRPPKCSLRCPHRCRINHFEBCCVKVDKCSLRCPHGECK
LHCAKTVTIDKENCCEKVPVCSNPLPDGSGHT"
19762..21336
/note="Gened exon scores (in order of location ranges) :
106.68 - GSCJ_ID dd_00725"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAS38632.1"
/db_xref="GI:42733688"
/translation="MDNNNLQPPYRPGSNNTCCDNDENGNDSTITQPKPAPSLPS
SQQSPITIIISQINHNKQLPKFEKHNNDISKMLCVLSEQISPKATNTPFDMLN
SLMLPPTPFLKLVMDHLKVIYIPYSVYNNVVSFVYDLSKNTALYDKYNTN
STYNNNNYNNVNNNSNNNADAKKKTITSATILNKENVELYNNNNNNNNNNNS
EILDKKSIKNNKVDGYITRILNLILACNSMDLSEINSINCAVNSOOFDPNNYIEK
SLDENQSOALQDOVCLARYPEVDVWRKDYLNHLETSQMSIEVLKQOOSKFPYVN
SSTIQPNKLSLVYIPYTYTKEQNSPFFLSNGQVGSATREPEIGLCTGEPFLK
TYNNFSLTVNGFEFSSCKQDEMLAIDSPKVTNTSSYFLTFRSSTNPLVSSSIGHL
TIIRNSKSIATLTIRGQKRSVQGNPILPLAHEFTSPFGHMCPPEIIDGEVNDQIVI
KINPEGNINQYIN"
complement(21961..24357)
/note="Gened exon scores (in order of location ranges) :
173.48 - GSCJ_ID dd_00724"
/codon_start=1
/product="similar to Arabidopsis thaliana (Mouse-ear
cress). Putative PREG-1-like negative regulator"
/protein_id="AAS38633.1"
/db_xref="GI:42733689"
/translation="MSKYSKSQYVSPKRPQNIPTTALPSSVASBLSLQKPPNAG
SHRKQRGSSSKNPNLPLOIDISTEKSSGKNNNNNNNNNNKRRKTSKNSVNSNNIN
SSSIVIKCEBQ000000000000000000000000DGFJLSQSEFNLFRNKLKXIDIS
```


SPSSSSSTSPSCLSSSSSSTFSOSIEAKENKODINNNNNNTNGNFEEEREKAKV
EQOIEERTFIESFEETKEVEPEPEPELPIISIGETIKETYSPLSYSEIIV
DMYNDKSSSTSSQSSSSSSSSSPSPKIRKNNNNNTTTTITTTTITTSKTATSS

Query Match 5.4%; Score 73.4; DB 3; Length 25769;
Best Local Similarity 49.4%; Pred. No. 0.00015;
Matches 217; Conservative 0; Mismatches 221; Indels 1; Gaps 1;

QY 438 CCGAACAAGATTAAATAAACAAGTAATCACTCGAAGCTATCATGTA 497
DB 2392 CGAACAACACGCTAAAAAATTTGTGAGAAAAAATAAATAAATAAATAA 2451
QY 498 TGTAAAGAAACATCTATTAAACAGATCCCTTAAATAAACAAGATATTCGAAG 557
DB 2452 TATGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2511
QY 558 AGACAAATTATGTTACAGTTTACAAATCTAAAGCGAATTAATATGAAAGTAAG 617
DB 2512 AAATTATTTATTAATATTAATTAATTAATTAATTAATTAATTAATTAATTA 2571
QY 618 CTATGACGTCAAGATTTTCTTTTCTGTTATTTGTTATTTGTTATTTATATCAT 677
DB 2572 ATAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2630
QY 678 TTTCTTCTTACATAGATGATTTTCTCCGATTTTAAATGACTATAAGTCAAT 737
DB 2631 TAATTTTAAAGATATTTGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 2690
QY 738 TTTATTAAGACAGCATGCTAGATCTCTGTTCAAAATCTTTCTGATTTTTTAAG 797
DB 2691 TTTAAATTTGGTGTGATTAATAATATCTTTTAAATTTTTCAGAAATATCTTT 2750
QY 798 AGCTAGTTGGCAACCTGTTCTTCAAGAAATTTGATTTTTCAAAAAATTAAGT 857
DB 2751 TTTTCTTTTCAACCCACCAATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 2810
QY 858 TATTTCTCTTATTAAT 876
DB 2811 TTAATTTTCTTTTATTTTAT 2829

RESULT 10
AF034389/c 713 bp DNA linear INV 04-FEB-1999
LOCUS AF034389
DEFINITION Plasmodium falciparum sexual stage antigen (s16) gene, promoter and partial cds.
ACCESSION AF034389
VERSION AF034389.1 GI:3098290
KEYWORDS
SOURCE Plasmodium falciparum (malaria parasite P. falciparum)
ORGANISM Plasmodium falciparum
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE 1 (bases 1 to 713)
AUTHORS Decherling,K.J., Kaan,A.M., Mbacham,W., Wirth,D.F., Eling,W., Koning,R.N. and Stunnenberg,H.G.
TITLE Isolation and functional characterization of two distinct sexual-stage-specific promoters of the human malaria parasite Plasmodium falciparum
JOURNAL Mol. Cell. Biol. 19 (2), 967-978 (1999)
MEDLINE 99108072
PUBMED 9891033
REFERENCE 2 (bases 1 to 713)
AUTHORS Decherling,K.J., Kaan,A.M. and Koning,R.N.H.
TITLE Direct Submision
JOURNAL Submitted (12-NOV-1997) Molecular Biology, University of Nijmegen, Toernooiveld 1, Nijmegen 6525 ED, The Netherlands
FEATURES
source 1..713
location/Qualifiers
1..713
/organism="Plasmodium falciparum"
/mol_type="genomic DNA"
/strain="NF54"
/db_xref="taxon:5833"

gene

/gene="s16"
1..536
/gene="s16"
537..>713
/gene="s16"
537..710
/gene="s16"
711..>713
/gene="s16"
/note="Pfs16"
/codon_start=1
/product="sexual stage antigen"
/protein_id="AA012581.1"
/db_xref="GI:3098291"
/translation="M"

ORIGIN

Query Match 5.4%; Score 71.2; DB 3; Length 713;
Best Local Similarity 45.8%; Pred. No. 0.00045;
Matches 282; Conservative 0; Mismatches 333; Indels 1; Gaps 1;

QY 444 CAAGATTTAAAAAATTAACCAAGTAATCACTCGAAGCTATCATGTAATGTTA 503
DB 681 CAAAAGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 622
QY 504 AAGAAACATCTATTAAACAGATCTCTTAAATAAACAAGATATTCGAAGACAA 563
DB 621 AATATATCTTATGATAAATAAAGCTATGTAATAATTAATTAATTAATTAAT 562
QY 564 ATATGTACAGTTTCAACAACATCTAAGACGCAATTAATTCGAAGTAAGCTATGA 623
DB 561 ATATATGTAATATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTT 503
QY 624 CGTTCAGATTTTCTTTTCTTCTGTTATTTGTTATTTGTTATTTTATTAATTTCT 683
DB 502 TTTTCTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 443
QY 684 CTCTTCAATTAAGCATTTTCTCCGATTTTAAATAAGACTATAAGTCAATTTTAA 743
DB 442 TTTTCAGTAAATAAGTATTAACGAAATATGTAATAAGAAATTTGGTAAAG 383
QY 744 TAAGACGACGATGCTGATTTCTGTTCAAAATCTTTGATTTTAAAGCTAG 803
DB 382 TAAATTTTAAATTAATAATAAATAAATAAATAAATAAATAAATAAATAAAG 323
QY 804 TTTGGCAACCTGTTCTTCAAGAAATTTGATTTTCAAAAAAATTAAGTTATTT 863
DB 322 TTTATTTATTTTAAATTAATTAATTTTAAATTTTAAATTTTAAATTTTAAATTT 263
QY 864 CTCTTATTAATAAGAAACCTTGAATAATTAAGCTTCCAGACTAGCCCTAGATGTT 923
DB 262 ATATTTTATTTTATTTTATTTATTAATAATAATAATAATAATAATAATAATTT 203
QY 924 TTCCCAATTAATTAACATCATGTAATTAATTTGGCAGCCCATTAATTAATTA 983
DB 202 TTTTCTTTTATTTTATTTTAAATTAATTTTATTTTATTTTCTACATTAATAATA 143
QY 984 CCGAACTGAATTCAGCGAACAATCTGACTATTTCTGATTAATTAATAAGGA 1043
DB 142 TAAATTTTATTTTAAAGGTATCATTAATGTTTATTTTAAATAAATAAATAATAGT 83
QY 1044 GAGAGAGAGAAAGAA 1059
DB 82 TATCTATATTAATAA 67

RESULT 11
DDU86962/c 3576 bp DNA linear INV 23-JUL-2001
LOCUS DDU86962
DEFINITION Dicycostejium discoideum Miga (miga) gene, complete cds.
ACCESSION U86962
VERSION U86962.1 GI:1841871
KEYWORDS

FEATURES	SOURCE
Location/Qualifiers	1. .3576
	/organism="Dicyostelium discoideum"
	/mol_type="genomic DNA"
	/strain="AX4"
	/sub strain="AK244"
	/db_xref="taxon:44689"
	/chromosome="4"
	/map="4.25"
	884. .3423
gene	/gene="miga"
	join(<884. .2899,2998. .>3423)
MRNA	/gene="miga"
	/product="Miga"
	join(884. .2899,2998. .3423)
CDS	/gene="miga"
	/function="essential for slug migration"
	/note="N-terminus of this protein is similar to other BTB domain proteins, such as the Mus musculus and human zinc finger proteins encoded by Genbank Accession Numbers U1556 and U16896, respectively"
	/codon_start=1
	/product="Miga"
	/protein_id="BAB47544.1"
	/db_xref="GI:1841872"
	/translation="MATIINNNNNNNNNNNNNINTEYFNNFGKSLDEPKKK TDKIKGGOIFPFRRIILANSSEFGRLILSDPKESQSYIELKQDPNNVPMVL FWDGIKIVSPENVILPLASDHYLIRLOHLCLIFLDONLHRENVTLIKSIDPH ETPERCITLIKSEFVGENLXKOTKINNNNNNCINENKHALTSSASSILANS TSASSSSSSNNNNYTTTTTTTTTTTASSNFKNLCVKG0000000000000000 0000000000PIYDSFLPASLFEYILDHDLVNVSEYTVLIVANIEKFEFENPT ITYNTITDOLKVSJDLMSQIRYLAMPYDQLITKDPILPEILFETILRQRGQ TSSPSSSTSSPSTNSPTKQTSRYG0000000000000000000000000000 PQ00PQ00PQ00PQ01000PQ01000PQ01000PQ01000PQ01000PQ010 PQ000 ISLMAQSKSLINNAATNTNNNNNNNNVYTNNTNNNTNNNNNNKQTSNNNDK LKRQYPRPQSLFEYSFDFKGIIFWISTDGENEKMSPHSTKIKITSSSIDGQ LVQIVELTPNAFMTKQDVPSAWMIDLGNKRVVPMYVYIRGLSYKSDSLRTDFOG TNEQWTVLKRHTNDPSLNKYKATHSWPTVCETAFRYFRILQTKGNRRNLVLVIG LEIYGELCETNPNP"
exon	<884. .2899
	/gene="miga"
	/number=1
misc_feature	951. .1289
	/gene="miga"
	/note="encodes BTB domain; similar to the Drosophila melanogaster BTB domain encoded by Genbank Accession Number U01333"
variation	1253^1254
	/gene="miga"
	/note="gene interruption in mutant"
intron	/replace="psori external plasmid, approximately 4 kb"
	2900. .2997
	/gene="miga"
	/number=1
exon	2998. .>3423
	/gene="miga"

[illegible]

Insert size: 149176; 6.5% error; agarose-fp
Quality coverage: 12.41x in Q20 bases; sum-of-contigs Quality
coverage: 12.54x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES

source
1. 149526
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone_1lb="CH211-117K16"
/clone_1lb="CHOR1-211"
1. 149526
/note="assembly_fragment:03115
clone_end:SP6
vector_side:right
clone_end:T7
vector_side:left"

ORIGIN

Query Match 5.4%; Score 70.8; DB 2; Length 149526;
Best Local Similarity 46.8%; Pred. No. 0.00047;
Matches 260; Conservative 0; Mismatches 292; Indels 4; Gaps 1;

QY 488 TATCATGTATGTTTAAAGAACATCTATTAAACACGATCTCTTAAACAAACAGCAT 547
DB 109063 TTTTAATATATGAACAAATATCTATATTTTACCTTTGCCCTTGTTAGCATTAAGCAT 109004
QY 548 ATTTCGAAAGAGCAATATTGTTAC-----GTTTACAAACATCTAAGCCGCAATTA 603
DB 109003 GTTTTAAAGTTTAAATTTTAAACATGTTTATATCAATATATGACTTAAAGAGA 108944
QY 604 TATCGAAAGGTATGCTATGACGTTGATGATTTTCTTTTCATTCCTGTTATTTGTTAT 663
DB 108943 CCAAAAGTACTAATTAGCATTTTAAATGTTTATTTTTCGATTTTATTTTATTTTAT 108884
QY 664 GTTTTATATATACATTTTCTCTCTTACATATAGATGATTTTCTTCGATTTTAAATG 723
DB 108883 AT 108824
QY 724 ACATATAAGCATTTTATATAGAGAGCAGCATGCTGATTTCTGTTCAAAATCTTT 783
DB 108823 TATATATTTTATTAATTTATATTAATAATATTATTAATAATTAATAATCAATTAC 108764
QY 784 CTGATTTTAAAGAGTATGTTGGCAACCGCTTTCTTCAAGAAATTTGATTTTTC 843
DB 108763 ATTATATCATATTTAACTCTTATATGTTTATATGTTATATATATATATATATAT 108704
QY 844 AAAAAAATTAAGTTTATTTCTCTTATTAATAATAGAAAACACTTAGAAAAATAGAGTGC 903
DB 108703 AGAAAAACAAAGTTATATTTATTTATTTCTTATATTTTAAATTAATAATTTT 108644
QY 904 CAGACTAGCCCTAGAAATGTTTCCCAATTAATTCAATCACTGTTATTAATTATTTGGCC 963
DB 108643 TATATATGCCCTTCATTAAGCATCTTATATTTTAAATAATATATATATTTTGGTC 108584
QY 964 AGCCCATTAATTAATTTAAACGGAATCTGAATCTGAGCAACCAATCGAGCTATTTC 1023
DB 108583 ACATATATATATTAAGAGAGACCAAAAGTAAATAATATGTTACTTATTTTATATCTGAA 108524
QY 1024 TCTAGATTAGTAAAAA 1039
DB 108523 GTTTTATATAGTTAAAA 108508

RESULT 13
PFMAL4P1/c PFMAL4P1 347582 bp DNA linear INV 29-JAN-2003
DEFINITION Plasmodium falciparum MAL4P1.
ACCESSION AL034557 AL844503

VERSION

AL034557.8 GI:23498126

KEYWORDS

Plasmodium falciparum 3D7

SOURCE

Plasmodium falciparum 3D7

ORGANISM

Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

REFERENCE

1

AUTHORS

Hall, N., Pain, A., Berriman, M., Churcher, C., Harris, B., Harris, D., Mungall, K., Bowman, S., Atkin, R., Baker, S., Barron, A., Brooks, K., Buckee, C.O., Burrows, C., Cherevach, I., Chillingworth, I., Chillingworth, T., Christodoulou, Z., Clark, L., Clark, R., Corton, C., Cronin, A., Davies, R., Davis, P., Dear, P., Dearden, F., Doggett, J., Felwell, T., Goble, A., Goodhead, I., Gilliam, R., Hamlin, N., Hance, Z., Harper, D., Hauser, H., Hornsby, T., Holroyd, S., Horrocks, P., Humphray, S., Jagels, K., James, K.D., Johnson, D., Kerhornou, A., Knights, A., Konfortov, B., Kyes, S., Larke, N., Lawson, D., Lennard, N., Line, A., Maddison, M., McLean, J., Mooney, P., Moulé, S., Murphy, L., Oliver, K., Ormond, D., Price, C., Quail, M.A., Rabinowitsch, E., Rajandream, M.A., Rutter, S., Sutherland, K.M., Sanders, M., Simmonds, M., Seeger, K., Sharp, S., Smith, R., Squares, R., Squares, S., Stevens, K., Taylor, K., Tivey, A., Unwin, L., Whitehead, S., Woodward, J., Suleston, J.B., Craig, A., Newbold, C. and Barrell, B.G.

TITLE

Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13

JOURNAL

Nature 419 (6906), 527-531 (2002)

MEDLINE

22255708

PUBMED

12368867

REFERENCE

2 (bases 1 to 347582)

AUTHORS

Hamlin, N., Pain, A., Berriman, B., Hall, N., Bowman, S., Churcher, C., Harris, B., Harris, D., Lawson, D., Quail, M. and Barrell, B.

TITLE

Submitted (24-FEB-1999) P.falciparum Genome Sequencing Consortium, The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK

COMMENT

On Oct 3, 2002 this sequence version replaced gi:5731897.
For more information about this sequence or the Malaria Project,
see http://www.sanger.ac.uk/Projects/P_falciparum.

FEATURES

Location/Qualifiers

1. 347582

/organism="Plasmodium falciparum 3D7"

/mol_type="genomic DNA"

/isolate="3D7"

/db_xref="taxon:36329"

/chromosome="4"

/clone="MAL4P1"

10384..32958

/note="rep20 Repeats"

join(35153..41725,42757..44124)

/gene="VAR"

/note="synonym: PPD0005w"

join(35153..41725,42757..44124)

/gene="VAR"

/note="similar to Plasmodium falciparum variant-specific surface protein var-2 SMALL-026033 (EMBL:U40609) (2664 aa)

fasta scores: E(): 7.8e-149, 43.1% id in 2872 aa"

/codon_start=1

/product="erythrocyte membrane protein 1 (PFEMP1)"

/protein_id="CAD49094.1"

/db_xref="GI:23498127"

/db_xref="GOA:O81220"

/db_xref="UniProt/TREMBL:O81220"

/translation="MWRVGGSGGGVVDGDDIDHSGAKHLSDSGKKVDQVKNAGDGT
GASDGNKYIDDLKDDIKAPNINPKLGTDDPCLVADVDYNNHNGGSGKRRYCTE
LSGKRFNPFSPTLGGOCITNSKMRSGSGGACAPYRRLHCHNLDSIETTSSTADSL
LLEVCMAAKVEGOSINTHYTKHSHNKSQPOLCVLARSFADIDIGIVGKDLFYGNT
YESAREKLKENTKEVPGKIHGGLSEAKKQYODDGNVYOLREDMWTANRETWKAI
TCEVSGNNVPRATGDEKNPSTLSKQCRCKDQKXGKPIKSGSNVNIPTVPDYVPOY
LWPEERNADEFCRLAKKLDKAIKKCRGNGEERKCDLNRDCKRTASGKVFPEEDF
CKDCQYSCAPFDWIDNQLLEFLKQKRYKTEITSGSGSGSGSGRKRRAATTNVEYE
KKFVTEKGLTKYKVNPNLEKLNDEVCCTKNNDIKDGNIDFKVNHSGSAKKGDGNLK
TFVPTKCGEACPMGCAKRVGGMKAKENCSQTKDYDDPKTTTIIILTGDTKSGDQK
KYKFCNGGNGGSGATPNATSRKGGKQDMEKICYIDENKEXKGTGSDAINCVL
ODGKHREKQVTSYNAPFMKWNHLDHSDVEMRRLNSCTINNAKSNQCKNKNCKE
CCGPEKWKQKKKEKEMBAIKHFRGQKQIIEOTGCDAGVTLAAVYLKLEFLNEDTEKS

EKGLDAEAKETIKHLRQMLEQAGVRLAAVGGPCTEGVAEQNTIMDKFLEBELKEAE
OCKNCPKPAQOQPGGARSADSPAGTGGDHPADDEDDDDDDDDDDDDDDDDDDDDDD
PCKNDLITSDPTKQVGDCEKRYGKGDHMKCGDITLVDPTVCMPREBERREED
YTIAMSEETKNIETODDLDAFIRTAALFTFSWQYKIKNGADQQLONGITIPBEFL
RSNYFTYGYDRLCLNTDISKTIVNDARAKADIKGRFSGDSSPGSTTPODMWQTY
GDIIMWGMICALHVTNTEKTKINDVSYDNQSONGNSPLEPARKPOFLMKI
EMGEERCAERKLEONIGKSCNGINPIQYCSNDRHPCNACDEKRYVTEKQEFEGQ
TKFEVADLENADQEKDYKTQPGSKOQNDYLDKCNKCSQMEGVLDIVDSKK
PGIYVAHKYSEKNCGLGAPKTPVNPAPRPOPAPPAIPATTPGVPCLVITLVS
SPNKKCACTLKCYGPKAPVPMKCIPTGNTNSGAAITDSEGDASRHRDAPSSGS
NQSICVPPRRKLYVPLTKAEBETTKSKQSGESKAGTSSSGSSASPGTISQ
GERSPGJSTPASTSSPNRDDLLKAEVESAAYETFLMHKYKMDKXELDEKKQ
OBSGVLGALDNGSNVSDBDKDPQKLEKISGIPBEFKQMFLLGDYDIDILVRGNT
SBSGNTNGNNNNIIVIASGDKODEMKIIOKADIENISLKOASVNPNGOQOON
SLTRETTLMKENHAPIWEGMICALTYKENDDEKIVDNEVEKEFGTTGTGSQYKE
KXYNTVAKLDENSDTEAKDTKTAAPSDNPTPLSHVLRPPRYIIEBGEFTCEKXK
KLAQIKVDCNDADYKCSGGESECKIDISITGVFADKCGCGRECKRYKATIER
KTIIEFQSKSAVYKQTKCKESGGGNGVCGTVKCTDAQFLERLQKRNNDGEG
KLEPNQSESTFKTKHCDPSSFKIDCRNGKSGDGTGKCDGITTIDAKELAKNIS
TPDVVVRSDNDPTNPEGDDLKYCEGKGIFKGRKEEMKRNCGGLDVGCLKKGND
KLDDKOIILIRALIKRWLEFLIEDYKIKHKISDCTNNGENICRDKCONKNCVGEW
ILCKEEMEKIKKHYLEKKEGDNMKSVRNPLEKEHPPENNAIKRCKLTQPS
FGLNGDKPSQNGHDAIDCMIKLEDKITSCLSSTSGQTETECQEHPLDDDEL
LBETENPVKQOPFCPCVKEKEKTEVDEKCGEDTSEPKTEDESKKKKDBSTPSP
EBPPAPPEAPPEKKEKVEBPQOPQPTTPOIVDKTAPALVTSTLAWSGVGAFTYFEL
KKKTSTIDLRVINIPKSDYDIPKLSPNRYIPTSGYRGKRYIYLEGDSGTDSG
TJHYSIDTSSSEYEEMDINDIYAPARPKYTLIEVLVPSKGNNTTASGNTTASGN
NTTASGNTTASGKTPSDTQNDIOSDGPSSKITDNMENTLKDEFISQYIOSQPKD
VNDPSSGDIPTFQNTNLTYPDEKPEFTSIHDNKLTYGBEYNDSTNSGNDLY
NKNKNLYSGQNNVSGIDPTSDNRGLTSGHDSYSGIDILNITLNGQNDIIDLVEKLY
KRNELFGTNHVTITNRPAPKARDPPLNOLLEPHTMLDHRNCKEMNKKEBLD
KKEWENETHTSGNTHPSDNTKLTNTDVSIOIDMDPKPINOFTNMIDVDPPTMDNM
EDDIYDVNDHDSITVDSNTMDVPSKVQIEMDVNTLTVKEKPIADVMDI"
45458 46601
/gene="PFD0010w"
/gene=".46601"
/note="VAR exon II, could be alternate exon II for upstream VAR gene PFD005w"
complement(join(48479..49537,49768..49821))
/gene="RIF"
/note="synonym: PFD0015c"
complement(join(48479..49537,49768..49821))
/gene="RIF"
/note="1 probable transmembrane helix predicted by TMHMM2.0 at aa 328-350
Similar to Plasmodium falciparum rifin pf00030C
SMALL:O96112 (EMBL:AE001367) (370 aa) fasta scores: E():
3..4e-65, 53..64 id in 375 aa"
/codon start=1
/product="Rifin"
/protein_id="CAD49095.1"
/db_xref="GI:23498128"
/translation="MKIHYINILFELPNTLILYNQRNHSNLTHTNRRSLCECELY
AATYADDPQMKEMVFKSQTOQRFHEDEVEKMKCKCKEIOKILIKQME
KELMDFAITQDIQSDALPTCYCEKSLADKVEKGLGAGVGLVAPVGLIGVAT
HWKPKALBAALAKAIAGETADIAAABAAGKAGREFTIKLKHGVGNFPGICDT
ISTSTGYTVTEFVNTIYKNGTCHUMSSINPTACITITBELSITKCGATGCHPP
LAIIRQIMIGLAEATEAKAABAANAQALTAITKELKE"
complement(join(52002..53297,54200..63307))
/gene="VAR"
/note="synonym: PFD0020c"
complement(join(52002..53297,54200..63307))
/gene="VAR"
/note="Similar to Plasmodium falciparum variant-specific
surface protein var-3 SMALL:026032 (EMBL:L40609) (3006 aa)
fasta scores: E(): 6..2e-103, 43..314 id in 3551 aa"
/codon start=1
/product="erythrocyte membrane protein 1 (PfEMP1)"
/protein_id="CAD49096.1"
/db_xref="GI:23498129"

/db_xref="GOA:081218"
/translation="WGTSSTPSVPKDVNBSHNARVLENGIEIYIEWEKRYNY
TSQNGDJSRAPFHGLKARLAVGIVPEPANSCDIDHPEYTNINNGVPPARPCLDRN
ONRFENAEACNSDKIRVTGKSSAGGACAPRRNMDCKNLEIDNTDITDGLK
NVLVTAKRENSIVAKHPKENSVECTALASPADIGDIVRCKDHFKNBEADAVGL
RAVFEKIDNLEKEISDYDNPYTKLREDMWTANRQWMAITCYIIPYVNYFEKT
SDDTLVFNDDCGYEGAPPTNLDIVPOPLRWPEMABEPRFINHKLTKIKDGRN
DKREIYVNGQYDCTKI EKGSCKREKCKCSKCYVDYDPELQKQNEFYQDQY
DEIETTYNKPITISNSNTKKEIYKERYBELKQYQYKPLQLNNGRYQOEKIE
EDALITPTNGHAYRSYDQPCDVCVECDGKCTCTQTDCKNRKIKLOKILISE
PIEIEVLVSDDKQVITEKLDIFGRGNNDENLOCMKCNKNSENKCMISMLY
ODPKENMLSYECFHSWAKNLLDITIMEHDLKNCINNTVCTDCTSKICEVLA
WIERKQDEMEKYLELANKODEFISYVKNLQDVPFRFQVWFLADQEGKMDPTE
DLTKRGVSVSAGTANSDOITBEYLHLKONALICGRNNISIKPTCPYBNPNCOT
NNGKLYVAVRIAEIMORPARKQLEKRGGEINLKADSQCKITIGKREKLNQICN
IDTSSVNSRNGNNGSPGLCKNDKFKI GTEVSGEHEKRRHPVYVMPREHEKIS
NLEKLDVSVIKNGNASHSLGDLVLALEKAKNLELYOQNSKNGVLDONDEKTC
RAMKSPADIGDIIIRKDMVYQNTATLQALVAKI PKIKONHDIKGLQYNGDTP
HKLAREDMANRHOVWRMKAICLINDKMCNGIPIEDYIPORLRMTWMAEWCXE
OSRLYGELEKQSGCKGOKCTEGVDGCKKCAADKYKDEINRREDMOTKIKGYCT
LYKATKRGVTTSNPKDREKVDYDLKQLPRKSNCTGVTAMPTNLTLSAAGTIHQ
ELGKTVGNTQEPFCDNKGKXYAFNHPREYEACTCOTROAQPIEKDNCNIGIT
LIDRSGGTGGIDGCPKTIQVPSNCRNBSKAKENKACVPPREKCVSLAKEGI
FKNKEDIRETFVSAALETYFAMRVYDNDKKEBELSGTIPENFRQMYTTPADY
RDIPEGTIDTSHDILDVSXNAKNLKEKNGQSKSVIIDDEKLADMMKEGHEIWE
GMICALTHEIDBEENKIKSYVSYDLKKTNGTTPLEKFAERPOFLWFTWMSDEFC
RERKEVEVEKCKCKDEGCKNPKTKNHCVSACDYEBIISTKQOYTOXKFP
IDKNGNEEYENYKQKAEADYLDIKCFPGTCDYWEKVNNSYWKPKYTNLSULK
KCECPQPPPARPAPVTOGACBIYVDILNGKASBPALICEKNGYKXGRPEMANSQI
RLOSARSDNPSPRPAPPGPQPSGTACAGARLETQSPQPPKPPAGNGVAPRAL
QPIARVDDDEEDDEEDDEESGSEEGEEDDEEDDEEDDEEDDEEDSHVNGCH
DUNQDEBGIIPDEPKROMYTFPGDYPFLGTSIDSGHAGTSELAKLIDSJFKNGG
KNPGLDSKMDNENNGEPIWKGMLCSLEAKNGKOTIKXNSYNNHNVNFSDRNBDL
ETPAKRPPLAMFTWMBEFCRBOQKQDILAKKPKSTCTNBEKCKKCSADKAYXE
WLOTKHEITEKQKITEKNDKSYTNDPTOKSPOAYQYLNKLEKICTSGNTSACEY
KCMKPPSSQNNNPNPASLDTPSDIKDICECTGASASRNSVRSDEBDPPRAP
ROSILARSADNPSPRPAPPGPQPSGTACAGARLETQSPQPPKPPAGNGVAPRAL
QPIARVDDDEEDDEEDDEESGSEEGEEDDEEDDEEDDEEDDEEDSHVNGCH
QEEBPPDETEVEVEETVAAPVYKPAICEIYKELPNDNRKDKCNLKYGGNNSLGMKCI
PSGDSSTTSSVNGDRSQHRRARAGATKSDASGICVPRRRRLRYVUKLTOMASQRT
QGGTSQIVGKTASQPNSHPTLPSBSNRDGLRDAFIOQAAVEFTFLMDRKXNLT
KPDATLGLPQIPLMGAINGVPSGDSNRDQKLEBIEIPBEFRQMFYTLGVDRI
LFGKNDIYIGMTSGASDKEMAKKEIKETIDKVPNSVSTPPTPAKTKSEDEKRT
WMEANGHEIWKMGYIALTYKNGEKKIYKDNSEVYKLPDEANKKKEKTEYQYKXKLE
ENSGAKPQPPSGSDNPTLTLNFTLS PVPYRIEEMGEPRRRKRLREYRKECR
GEYPERKICGEGDHDCTNGELKHTNMSADBDPCDHQCKRRTMIDIKREBYKOK
DKYQGEIDLKNGNSNGNCKEIKGHTSASPELKHCHCDQGSBDDTDSBEDKK
NNKIDENPLETFNPSTYCEFCPSKNKVCNSGRQTRKDPCTPNEKQKSEVFN
NGNSSTEITVEMIDRBLIKNYSKILSESGNSDSLFTKSTSLPFSVNRDQWECRYKD
EKTDICKLNFNDKIDLNOYTFKVFLELODFEGYIILKRKIIEOCKENGGSET
NENSKNDACVKGWVAQKTEMNQIKOHYNNKKEYGNGVDMSHKXNGVEKONNELKRV

```

Db      335424 TTTTCAGTATATTAAGGAAATGTAGAAATGGAAGAAAGATTGGATG 335365
Qy      744 TAAGAGACCGAGTCGTAGATTCTGCTCAAAAATCTTCTGATTTTAAAGAGTAG 803
Db      335364 TAAATTTTAAATATATAATATATAATATATATATATATATATATATATATATATAT 335305
Qy      804 TTGGCAACCTGTTCTTCAAGAAATTTGATTTTTCAGAAAAAATTAAGTTATTT 863
Db      335304 TTTATTTTATATATATATATATATATATATATATATATATATATATATATATAT 335245
Qy      864 CTCTTATATATATATATATATATATATATATATATATATATATATATATATAT 923
Db      335244 ATATTTTATATATATATATATATATATATATATATATATATATATATAT 335185
Qy      924 TTCCCAATATATATATATATATATATATATATATATATATATATATATATAT 983
Db      335184 TTTTATTTTATATATATATATATATATATATATATATATATATATATATATATAT 335125
Qy      984 CCGAAATGAAATCGAGCGAAACCAATCTGAGTATTTCTTATGATTAAGGA 1043
Db      335124 TAAATTTTATATATATATATATATATATATATATATATATATATATATAT 335065
Qy      1044 GAGAGAGAGAGAGAA 1059
Db      335064 TATCTATATATATATATATATATATATATATATATATATATATATAT

```

RESULT 14
CQ422641/c
LOCUS CQ422641 883 bp DNA linear PAT 28-JAN-2004

DEFINITION Sequence 7675 from Patent WO0151628.
ACCESSION CQ422641
VERSION CQ422641.1 GI:41374870
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1
AUTHORS Lillie, J., Xu, Y., Wang, Y. and Steinmann, K.
TITLE Novel genes, compositions, kits, and methods for identification, assessment, prevention, and therapy of breast cancer
JOURNAL Patent: WO 0151628-A 7675 19-JUL-2001;
Milenium Pharmaceuticals, Inc. (US)
FEATURES
Source 1..883
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 5.4%; Score 70.2; DB 6; Length 883;
Best Local Similarity 39.2%; Pred. No. 0.0007;
Matches 237; Conservative 0; Mismatches 367; Indels 0; Gaps 0;

```

Qy      419 TGTCCACTATTGGCACTACCCAGAACAGATTTAAAAAATAACCAAGTAAGTAATCCA 478
Db      873 TGTNNANNNNNACCTTAACCTTTAAATATNTNNNANTNNANNAAAAAAATATTTAAANN 814
Qy      479 CTGGAAGCTATGATGTAATGTTTAAAGAAACATCTATTAAACACGATCCTTTAAA 538
Db      813 TAAAAATTTTNNAAAAAATAATNTNANNATNNTNNTTNNAAAAAANNAAAAA 754
Qy      539 AACAGCATATTTGAAAGAGACAAATATGTTTACAGTTTACAAACATCTAAGAGCACA 598
Db      753 TTAANNNTTTTNTAANTATTAACCAAAATTTTAAAAAATTTTAAAAAANTT 694
Qy      599 AATTATATCGAAAGTAAGTATGACGTTCAGATTTTCTTTTCAATTTGTTATTTTG 658
Db      693 AATTAANTTNAATTTTATATNAAAAAATTTTAAATTTTAAACAAANTTTT 634
Qy      659 TTAATGTTTATATACATTTTCTCTTACAAATAGAGATTTTCTTCGATTTTATA 718

```

```

Db      633 TTTNTTNNATTAATAAATTTTAAATTAATAAANNANNTTTTATATNANATATTA 574
Qy      719 AAATGACTATPAAGTATTTTATATPAAGACGAGATGCTAGATTCTGTTCAAAA 778
Db      573 AAAAAAANNTTTTATNTTATNTTATTAATAAATAATTTATTTNTTTTCTNAAAAA 514
Qy      779 TCTTTCTGATTTTTPAAGAGCTAGTTGGCAACCTGTTCTTTCAAGAAATTTGATT 838
Db      513 AAAAAAATTTTNNAAANNTTTTAAAACTTNNATNNAAANNAATTTTNTTNTT 454
Qy      839 TTTTCAAAAAATTTAGTTATTTCTCTTATTAATTAAGAAACCTTAGAAAAATAGA 898
Db      453 NNANNTTAATAAATTTATTTTNTTCTATTAATAAATAAANNAATTAATAAT 394
Qy      899 GTTGCAGACTAGCCCTAGATGTTTCCCAATPAATTAACAATCAGTGTATATATAT 958
Db      393 TTTNAAAAATATATTTTATATNTATNTNANANNAATTTTAAANTNANNAATAAT 334
Qy      959 TGCCAGCCCATTAATTTATTAACCGAAACTGAATGAGGAAACCAATCTGAGCT 1018
Db      333 NTTAAATTAATTAATAAATTAATTTAAACATTTNTTAATTTATTAATTAATAA 274
Qy      1019 ATTT 1022
Db      273 NTTT 270

```

RESULT 15
BV119878/c
LOCUS BV119878 241 bp DNA linear STS 19-MAR-2004

DEFINITION PZA01377 CML247 Zea mays CML247 Zea mays STS genomic, sequence tagged site.
ACCESSION BV119878
VERSION BV119878.1 GI:45589251
KEYWORDS STS.
SOURCE Zea mays
ORGANISM Zea mays

REFERENCE 1
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
JOURNAL 1 (bases 1 to 241)
Mcmullen, M.D., Vroh Bi, I., Schroeder, S.S. and Gaut, B.S.
MPZ-OCI Joint SNP Discovery
Unpublished (2003)
COMMENT

Contact: Brandon S. Gaut
Dept. Ecology and Evolutionary Biology
U.C. Irvine
321 Steinhilber Hall, Irvine, CA 92697-2525, USA
Tel: (949) 824-2564
Fax: (949) 824-2181
Email: bgaut@uci.edu

Primer A: TTCTACGTCGAATGATCAAGA
Primer B: AGGAACCAAGCTTATGAGAGA
STS size: 241
Protocol:

PCR amplification of genomic DNA
Template: 50 ng
Primer: each 0.5 uM
dNTPs: each 200 uM
Tag Polymerase: RedTaq (Sigma)
Total Vol: 10 uL

Amplicon sequencing
ABI protocol - using d-Rhodamine terminator cycle
Sequencing ready reaction with ampliTaq DNA polymerase FS
Sequence ran on ABI 3700 sequencer.

Buffer:
Genomic DNA amplification
RedTaq (Sigma)
Sequencing buffer
d-Rhodamine kit (ABI)


```

XX Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;
SQ
Query Match      100.0%; Score 1311; DB 2; Length 1394;
Best Local Similarity 100.0%; Pred. No. 1e-269;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCATGGTGTCTCTATGAAAAAGATGAGTCAATGTGTCTATCTCGTTTCTTAAAGGATCC 60
Db 1 CCATGGTGTCTCTATGAAAAAGATGAGTCAATGTGTCTATCTCGTTTCTTAAAGGATCC 60

Qy 61 CTTCCTCTGCTTATTAATGACTGAAATCGGGGTTTAACAAAAATTCCAGGGTGCATGAT 120
Db 61 CTTCCTCTGCTTATTAATGACTGAAATCGGGGTTTAACAAAAATTCCAGGGTGCATGAT 120

Qy 121 CTTCATGTTCCACTTCTCCACCTCGGCTTGCACTTCTTGGAATGTCGGTGTCCCAT 180
Db 121 CTTCATGTTCCACTTCTCCACCTCGGCTTGCACTTCTTGGAATGTCGGTGTCCCAT 180

Qy 181 CTGACCGAGGCCCATGACACCTTTCGGGACACCATCAAGGGCTTTGATGGCCCA 240
Db 181 CTGACCGAGGCCCATGACACCTTTCGGGACACCATCAAGGGCTTTGATGGCCCA 240

Qy 241 CGAGACGTATCGGCTGTGTGATTCAGGGGATATATGTCGCCACATGTCACCTATA 300
Db 241 CGAGACGTATCGGCTGTGTGATTCAGGGGATATATGTCGCCACATGTCACCTATA 300

Qy 301 TTATATCTCTTATGATATTAATTTTGGAAAAATAACAACCTTATCTTTGTTGTA 360
Db 301 TTATATCTCTTATGATATTAATTTTGGAAAAATAACAACCTTATCTTTGTTGTA 360

Qy 361 GGGCCTCAGCATGATTTTGCCTTAAAGGCCAGAAAATGCGAGGACCGCATCTCAATG 420
Db 361 GGGCCTCAGCATGATTTTGCCTTAAAGGCCAGAAAATGCGAGGACCGCATCTCAATG 420

Qy 421 TCCACTTATGCGACTACCCAGAAACAAGTTTAAAAAATACCAAGTAACTATATCACT 480
Db 421 TCCACTTATGCGACTACCCAGAAACAAGTTTAAAAAATACCAAGTAACTATATCACT 480

Qy 481 CGAAGCTATCATGATATGTTTAAAGAAACATCTATTTAAACAGACGCTCTTAAAAA 540
Db 481 CGAAGCTATCATGATATGTTTAAAGAAACATCTATTTAAACAGACGCTCTTAAAAA 540

Qy 541 CAAGCATATTTGAAAAAGACAAATATGTTTACAGTTTAAACAATCTTAAGAGCGACAA 600
Db 541 CAAGCATATTTGAAAAAGACAAATATGTTTACAGTTTAAACAATCTTAAGAGCGACAA 600

Qy 601 TTATATCGAAAGGTAAAGCTATGACGTTTCAGATTTTCTTTTCAATCTTGTATTTGTT 660
Db 601 TTATATCGAAAGGTAAAGCTATGACGTTTCAGATTTTCTTTTCAATCTTGTATTTGTT 660

Qy 661 ATGTTTTTATATACATTTTCTCTTACAAATAGATGATTTCTTCGATTTTAAAA 720
Db 661 ATGTTTTTATATACATTTTCTCTTACAAATAGATGATTTCTTCGATTTTAAAA 720

Qy 721 ATGACTATAAAGTCATTTTATATAAAGACGACGATGCTAGATTTCTGTTCAAAAATC 780
Db 721 ATGACTATAAAGTCATTTTATATAAAGACGACGATGCTAGATTTCTGTTCAAAAATC 780

Qy 781 TTTCTGATTTTAAAAAGAGCTAGTTGGCAACCTGTTCTTCAAAAGATTTGATTTT 840
Db 781 TTTCTGATTTTAAAAAGAGCTAGTTGGCAACCTGTTCTTCAAAAGATTTGATTTT 840

Qy 841 TTTCTGATTTTAAAAAGAGCTAGTTGGCAACCTGTTCTTCAAAAGATTTGATTTT 840
Db 841 TTTCTGATTTTAAAAAGAGCTAGTTGGCAACCTGTTCTTCAAAAGATTTGATTTT 840

Qy 901 TGCCAGACTAGCCCTAGAAATGTTTCCAAATAATTAACAATCACTGTGATATATTTTG 960
Db 901 TGCCAGACTAGCCCTAGAAATGTTTCCAAATAATTAACAATCACTGTGATATATTTTG 960

Qy 961 GCCAGGCCCTAAATTTTAAACCGAAATCGAGCGAAACCAATCTGAGCTAT 1020
Db 961 GCCAGGCCCTAAATTTTAAACCGAAATCGAGCGAAACCAATCTGAGCTAT 1020

```

```

Db 961 GCCAGGCCCTAAATTTTAAACCGAAATCGAGCGAAACCAATCTGAGCTAT 1020
Qy 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGGAGAAATCACTTTTAAGTATGTCCTCC 1080
Db 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGGAGAAATCACTTTTAAGTATGTCCTCC 1080

Qy 1081 TGAGATGTGCGGTTTGGCAACGATGACCGCTAATCAATAGCTCATAGGTGCTTACGTA 1140
Db 1081 TGAGATGTGCGGTTTGGCAACGATGACCGCTAATCAATAGCTCATAGGTGCTTACGTA 1140

Qy 1141 GGTGGGCGAGCTCTGTGTGATCTCAATGGCATATCTACATCTGTTTAAACCGTTGTC 1200
Db 1141 GGTGGGCGAGCTCTGTGTGATCTCAATGGCATATCTACATCTGTTTAAACCGTTGTC 1200

Qy 1201 TTGTTCCATTCGTCGAAGCCTTGCTTATCTGAAACCAAGAGATACCTACTCCCAACAT 1260
Db 1201 TTGTTCCATTCGTCGAAGCCTTGCTTATCTGAAACCAAGAGATACCTACTCCCAACAT 1260

Qy 1261 CCATCTTACTCATGCACTTCCATGCAACGCAACATATGTTTCTGAAAC 1311
Db 1261 CCATCTTACTCATGCACTTCCATGCAACGCAACATATGTTTCTGAAAC 1311

```

```

RESULT 2
AAK07409
ID AAK07409 strand; DNA; 1394 BP.
XX
AC AAK07409;
XX
DT 08-JUN-1999 (first entry)
XX
DE Zea mays M45 male tissue-preferred regulatory region.
XX
KM Ms45; male; tissue-preferred; regulatory region; plant cells;
XX plant tissue; differentiated; hybrid seed; fertility; ss.
XX
OS Zea mays.
XX
PN MO9859061-A1.
XX
PD 30-DEC-1998.
XX
PF 19-JUN-1998; 98WO-US012895.
XX
PR 23-JUN-1997; 97US-00880499.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albreten MC, Fox TW, Garnaat CW, Huffman GA, Kendall TL;
XX MPI; 1999-105628/09.
XX
PT New nucleic acid encoding a Ms45 male tissue-preferred regulatory region.
XX - useful in mediating plant fertility, especially hybrid seed production.
XX
XX Claim 3; Page 23-24; 39pp; English.
XX
The sequence is that encoding an Ms45 male tissue-preferred regulatory
CC region. It may be used in the construction of a vector for a method of
CC producing exogenous genes in a male tissue-preferred manner, which is
CC useful in restoring or conferring fertility, such as in hybrid seed
CC production. In conferring fertility, a monocot/dicot plant is transformed
CC with the exogenous nucleotide sequence (a male sterility gene, preferably
CC Ms45), which encodes a product selected from auxins, rolb and dipteris
CC toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
CC and infertile plants
CC
SQ Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;
XX
Query Match      100.0%; Score 1311; DB 2; Length 1394;
Best Local Similarity 100.0%; Pred. No. 1e-269;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```



```

Oy 1 CCAATGCTCTCTATGAAAAAGATGATGATGCTATATCCGTTTCTTAAAGGATCC 60
Db 1 CCAATGCTCTCTATGAAAAAGATGATGATGCTATATCCGTTTCTTAAAGGATCC 60
Oy 61 CTTCTTCTGCTTATTAAGTAAATCGGGGTTTACAAAAAATTCCACGGGTGCATGAT 120
Db 61 CTTCTTCTGCTTATTAAGTAAATCGGGGTTTACAAAAAATTCCACGGGTGCATGAT 120
Oy 121 CTGCAATGTCACCTTCTCCCACTCGGGTTCACATTTCTTGATGTCGGTGGTCCAT 180
Db 121 CTGCAATGTCACCTTCTCCCACTCGGGTTCACATTTCTTGATGTCGGTGGTCCAT 180
Oy 181 CTGACCGAGGCCCATCGACACCTTTGGGACACCCATCAAGGGCCCTTGGATGGCCCA 240
Db 181 CTGACCGAGGCCCATCGACACCTTTGGGACACCCATCAAGGGCCCTTGGATGGCCCA 240
Oy 241 CGAAGCGTATCGGGTCGTGATTCAGGGGATATATGTCCTCCCAATGTCACCTATA 300
Db 241 CGAAGCGTATCGGGTCGTGATTCAGGGGATATATGTCCTCCCAATGTCACCTATA 300
Oy 301 TTATATCTTCTTATGATTTATTTATTTTGGAAAAATACAACTTATCTTTGTGTA 360
Db 301 TTATATCTTCTTATGATTTATTTATTTTGGAAAAATACAACTTATCTTTGTGTA 360
Oy 361 GGGCCTCAGCATGATTTGCTTAAAGGGCCAGAAATGAGGAGCCAGCATGTCATGTC 420
Db 361 GGGCCTCAGCATGATTTGCTTAAAGGGCCAGAAATGAGGAGCCAGCATGTCATGTC 420
Oy 421 TCCATATTTGGCACTACCCAGAACAGATTTAAAAAATAACAAAGTAACTATCCACT 480
Db 421 TCCATATTTGGCACTACCCAGAACAGATTTAAAAAATAACAAAGTAACTATCCACT 480
Oy 481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAACCAAGTCTCTTAAAAAA 540
Db 481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAACCAAGTCTCTTAAAAAA 540
Oy 541 CAAGCATATTTGCAAGAGAGCAAAATTAATGTTACAGTTTCAAAATATAGAGCGACAA 600
Db 541 CAAGCATATTTGCAAGAGAGCAAAATTAATGTTACAGTTTCAAAATATAGAGCGACAA 600
Oy 601 TTATATCGAAAGGTAAAGCTATGACGTTCAAGATTTTCTTTTCAATCTTGTATTTTGT 660
Db 601 TTATATCGAAAGGTAAAGCTATGACGTTCAAGATTTTCTTTTCAATCTTGTATTTTGT 660
Oy 661 ATTGTTTTTATATACATTTTCTCTCTTCAATAGAGTATTTTCTTCCGATTTTATAAA 720
Db 661 ATTGTTTTTATATACATTTTCTCTCTTCAATAGAGTATTTTCTTCCGATTTTATAAA 720
Oy 721 ATGACTATTAAGTCAATTTTATATATAGAGACGCAATGTCGTAGATTCGTCAAAAAATC 780
Db 721 ATGACTATTAAGTCAATTTTATATATAGAGACGCAATGTCGTAGATTCGTCAAAAAATC 780
Oy 781 TTTCTGATTTTTTAAAGAGCTAGTTTGGCAACCTGTTTCTTCAAGAAATTTGATTTT 840
Db 781 TTTCTGATTTTTTAAAGAGCTAGTTTGGCAACCTGTTTCTTCAAGAAATTTGATTTT 840
Oy 841 TTCAAAAAAATTAAGTTATTTTCTCTTAAATAGAAAAACATTAAGAAAAATAGAGT 900
Db 841 TTCAAAAAAATTAAGTTATTTTCTCTTAAATAGAAAAACATTAAGAAAAATAGAGT 900
Oy 901 TGGCAGACTAGCCCTGAATGTTTCCCATTAATTAATCACTGCTGATTAATTTTGG 960
Db 901 TGGCAGACTAGCCCTGAATGTTTCCCATTAATTAATCACTGCTGATTAATTTTGG 960
Oy 961 GCCAGGCCCAATTAATTTTAAACCGAAATCGAAATCGAGAAACCAATCTGAGCTAT 1020
Db 961 GCCAGGCCCAATTAATTTTAAACCGAAATCGAAATCGAGAAACCAATCTGAGCTAT 1020
Oy 1021 TTCTCTAGATTTAGTAAAAAGGAGAGAGAGAGAGAAATCACTGTTTAAGTCTATGCC 1080
Db 1021 TTCTCTAGATTTAGTAAAAAGGAGAGAGAGAGAGAAATCACTGTTTAAGTCTATGCC 1080
Oy 1081 TGAGATGTGGGTTTGGCAACGATGACCATGATATGCTATGATGCTAGCTAC 1140

```

```

Db 1081 TGAGATGTGGGTTTGGCAACGATGACCATGATATGCTATGATGCTAGCTAC 1140
Oy 1141 GGTTCGGAGCTCTCGTGTATCTCATGTCATGAGCATATGATGTTTGAACCGTTGTC 1200
Db 1141 GGTTCGGAGCTCTCGTGTATCTCATGTCATGAGCATATGATGTTTGAACCGTTGTC 1200
Oy 1201 TTGTTCCATGTCGTCAGAGCCCTTGGCTTATTTGGAACCAAGAGATACTTCCAAACAT 1260
Db 1201 TTGTTCCATGTCGTCAGAGCCCTTGGCTTATTTGGAACCAAGAGATACTTCCAAACAT 1260
Oy 1261 CCATCTTACTCATGCAACTTCCATGCAACACGACATATGTTTCTGAAAC 1311
Db 1261 CCATCTTACTCATGCAACTTCCATGCAACACGACATATGTTTCTGAAAC 1311

RESULT 3
AAH76332
ID AAH76332 standard; DNA; 1394 BP.
XX
AC AAH76332;
XX
DT 29-OCT-2001 (first entry)
XX
DE Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.
XX
KW Ms45; male tissue; regulatory region; transcription; male fertility;
KW hybrid seed; da.
XX
OS Zea mays.
XX
PN WO200160997-A2.
XX
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
DR WPI; 2001-514772/56.
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the Ms45 gene useful for
PT mediating fertility in a male plant.
XX
PS Claim 4; Page 46; 50pp; English.
XX
CC The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the Ms45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a nucleic acid sequence encoding an Ms45 male
CC tissue preferred regulatory region from Z. mays
XX
SQ Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;

Query Match 100.0%; Score 1311; DB 5; Length 1394;
Best Local Similarity 100.0%; Pred. No. 1e-269;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CCAATGCTCTCTATGAAAAAGATGATGATGCTATATCCGTTTCTTAAAGGATCC 60
Db 1 CCAATGCTCTCTATGAAAAAGATGATGATGCTATATCCGTTTCTTAAAGGATCC 60
Oy 61 CTTCTTCTGCTTATTAAGTAAATCGGGGTTTACAAAAAATTCCACGGGTGCATGAT 120

```

```

Db      61 CTCTCTGCTTAACTGACTGAATCGGGGTTACAAAAAATTCCACGGGTGCATGAT 120
QY      122 CTCATGTTGCACCTTCCACCTCGGCTTGACATTTCTGGATGCGGTGGTCCCAT 180
Db      121 CTCATGTTGCACCTTCCACCTCGGCTTGACATTTCTGGATGCGGTGGTCCCAT 180
QY      181 CTGACCGAGGCCCATGACACACCTTTCGGGACACCATCAAGGCGCTTTCGATGGCCCA 240
Db      181 CTGACCGAGGCCCATGACACACCTTTCGGGACACCATCAAGGCGCTTTCGATGGCCCA 240
QY      241 CGAGAGGTATCGGGTGTGTGATCCAGGGGATATATGTCCCCCAACATGTCACTTAA 300
Db      241 CGAGAGGTATCGGGTGTGTGATCCAGGGGATATATGTCCCCCAACATGTCACTTAA 300
QY      301 TTATATTTCTTGAATTTATTTAATTTTGAATAAATAAATACTTATCTTTTGTGTA 360
Db      301 TTATATTTCTTGAATTTATTTAATTTTGAATAAATAAATACTTATCTTTTGTGTA 360
QY      361 GGGCTTCAGCATGATTTTTCGCTTAAAGGCCAGAAAATGCGAGGACCGCATGTCTAGTG 420
Db      361 GGGCTTCAGCATGATTTTTCGCTTAAAGGCCAGAAAATGCGAGGACCGCATGTCTAGTG 420
QY      421 TCCACTATTGGCACTACCCAGAAAGATTTAAATAAATAAATACTTATCTTCACT 480
Db      421 TCCACTATTGGCACTACCCAGAAAGATTTAAATAAATAAATACTTATCTTCACT 480
QY      481 CGAAGCTATCATGATATGTTTAAAGAAACATCTATTTAAACACGATCCTCTTAAATA 540
Db      481 CGAAGCTATCATGATATGTTTAAAGAAACATCTATTTAAACACGATCCTCTTAAATA 540
QY      541 CAAGCATATTTGAAAAGAGCAAAATATGTTACAGTTTACAAACATCTAAGACGACAA 600
Db      541 CAAGCATATTTGAAAAGAGCAAAATATGTTACAGTTTACAAACATCTAAGACGACAA 600
QY      601 TTATATCGAAAGGTAAAGCTATGACGTTCAAGATTTTCTTTTATCTTGTATTTGTT 660
Db      601 TTATATCGAAAGGTAAAGCTATGACGTTCAAGATTTTCTTTTATCTTGTATTTGTT 660
QY      661 ATTGTTTTATATACATTTTCTTCTTCAATAGAGTATTTCTTCCGATTTTATTA 720
Db      661 ATTGTTTTATATACATTTTCTTCTTCAATAGAGTATTTCTTCCGATTTTATTA 720
QY      721 ATGACTATTAAGTCAATTTTATTAAGAGACGCGATGCTAGATTTCTCGTTCAAAATC 780
Db      721 ATGACTATTAAGTCAATTTTATTAAGAGACGCGATGCTAGATTTCTCGTTCAAAATC 780
QY      781 TTTCTGATTTTAAAGAGTATTTGGCAACCTGTTCTTCAAGAAATTTGATTTT 840
Db      781 TTTCTGATTTTAAAGAGTATTTGGCAACCTGTTCTTCAAGAAATTTGATTTT 840
QY      841 TTCAAAAAAAATAGTTTATTTCTCTTATTAATAAGAAAACTTAGAAAAATAGAGT 900
Db      841 TTCAAAAAAAATAGTTTATTTCTCTTATTAATAAGAAAACTTAGAAAAATAGAGT 900
QY      901 TGCAGACTAGCCCTAGAAATGTTTTCCCAATAAATTAACATCCTGTGTAATTAATTTG 960
Db      901 TGCAGACTAGCCCTAGAAATGTTTTCCCAATAAATTAACATCCTGTGTAATTAATTTG 960
QY      961 GCGAGCCCATTAATTTATTAACCGAAACTGAATTCGAGCGGAACCAATCTGAGCTAT 1020
Db      961 GCGAGCCCATTAATTTATTAACCGAAACTGAATTCGAGCGGAACCAATCTGAGCTAT 1020
QY      1021 TTCTCTAGATTAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Db      1021 TTCTCTAGATTAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
QY      1081 TGAAGATGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCTTACGTCA 1140
Db      1081 TGAAGATGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCTTACGTCA 1140
QY      1141 GGTTCGGGAGCTTCTGTCATCTCACATGCGATCTACTACATGCTTGTCAACGTTTCGTC 1200
Db      1141 GGTTCGGGAGCTTCTGTCATCTCACATGCGATCTACTACATGCTTGTCAACGTTTCGTC 1200

```

```

QY      1201 TTGTTCCATGTCGCAAGCCTTGCTTATTTCTGAACCAAGAGATPACTACTCCAAACAT 1260
Db      1201 TTGTTCCATGTCGCAAGCCTTGCTTATTTCTGAACCAAGAGATPACTACTCCAAACAT 1260
QY      1261 CCATCTTACTCATGCAACTTTCATGCAACACGACATATGTTTCTGTAAC 1311
Db      1261 CCATCTTACTCATGCAACTTTCATGCAACACGACATATGTTTCTGTAAC 1311

```

RESULT 4

AAH76333 standard; DNA; 1394 BP.

AAH76333;

29-OCT-2001 (first entry)

Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.

Ms45; male tissue; regulatory region; transcription; male fertility; hybrid seed; ds.

Zea mays.

MO200160997-A2.

23-AUG-2001.

13-FEB-2001; 2001WO-US004527.

15-FEB-2000; 2000US-00504487.

(PION-) PIONEER HI-BRED INT INC.

Albertain MC, Fox TW, Garmaat CW, Huffman G, Kendall TL;

WPI; 2001-514772/56.

A male tissue-preferred regulatory region comprising nucleotide sequences essential for initiating transcription of the Ms45 gene useful for mediating fertility in a male plant.

Claim 4; Page 47; 50bp; English.

The invention provides a male tissue-preferred regulatory region (i) comprising nucleotide sequences essential for initiating transcription of the Ms45 gene. A method of mediating male fertility in a plant is provided that involves introducing an expression vector comprising a promoter operably linked to (i) into a plant where the exogenous gene impacts male fertility of the plant and (ii) controls expression of the exogenous gene. A method of producing hybrid seeds is also provided. The present sequence represents a nucleic acid sequence encoding an Ms45 male tissue preferred regulatory region from Z. mays

Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;

Query Match 100.0%; Score 1311; DB 5; Length 1394;

Best Local Similarity 100.0%; Pred. No. 1e-269; Mismatches 0; Indels 0; Gaps 0;

Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CCATGATGTCCTATGAAAAAGATGATACATGTCATATCCGTTTCTTAGGGTCC 60

1 CCATGATGTCCTATGAAAAAGATGATACATGTCATATCCGTTTCTTAGGGTCC 60

61 CTCTCTTGCCTTATTACTGACTGAATGGGGTTACAAAAAATCTTCACGGGTGCATGAT 120

61 CTCTCTTGCCTTATTACTGACTGAATGGGGTTACAAAAAATCTTCACGGGTGCATGAT 120

121 CTCATGTTCCACTTCTCCACCTGCGTTGACATTTCTGGAGTGGGTGCCAT 180

121 CTCATGTTCCACTTCTCCACCTGCGTTGACATTTCTGGAGTGGGTGCCAT 180

QY 181 CTGACGAGGCGCCATGAGACACCTTTGCGGACCAACCAAGGGCTTTGCGATGGCCCA 240
 DB 181 CTTGACGAGGCGCCATGAGACACCTTTGCGGACCAACCAAGGGCTTTGCGATGGCCCA 240
 QY 241 CGAGACGTATCGGGTCGTGGTATCGAGGGATATATGTCCCAACATCGTCACTATA 300
 DB 241 CGAGACGTATCGGGTCGTGGTATCGAGGGATATATGTCCCAACATCGTCACTATA 300
 QY 301 TTATATCTTTTATGATATTTATTTAAATTTTGGAAAAATTAACAACTTATCTTTGGTGA 360
 DB 301 TTATATCTTTTATGATATTTATTTAAATTTTGGAAAAATTAACAACTTATCTTTGGTGA 360
 QY 361 GGGCCCTCAGCATGATTTTGGCTTAAAGGCGCAGAAATGCGAGACCAAGCATGTCTAGTG 420
 DB 361 GGGCCCTCAGCATGATTTTGGCTTAAAGGCGCAGAAATGCGAGACCAAGCATGTCTAGTG 420
 QY 421 TCCACTATTTGGCACTACCCAGAACAGATTTAAAAAATAACCAAGTAACTATCCACT 480
 DB 421 TCCACTATTTGGCACTACCCAGAACAGATTTAAAAAATAACCAAGTAACTATCCACT 480
 QY 481 CGAAAGCTATCATGTATGTTTAAAGAAACATCTATTTAAACCAAGTCTTTAAAAAA 540
 DB 481 CGAAAGCTATCATGTATGTTTAAAGAAACATCTATTTAAACCAAGTCTTTAAAAAA 540
 QY 541 CAAGCATATTTGCAAGAGACAAATTTATGTTACAGTTTACAAACATCTAGAGCGACAA 600
 DB 541 CAAGCATATTTGCAAGAGACAAATTTATGTTACAGTTTACAAACATCTAGAGCGACAA 600
 QY 601 TTAATTCGAAAGGTAGCTATGACGTTCAGATTTTCTTTTCACTCTGTGTAATTTGTT 660
 DB 601 TTAATTCGAAAGGTAGCTATGACGTTCAGATTTTCTTTTCACTCTGTGTAATTTGTT 660
 QY 661 ATTGTTTTTATATACATTTTCTTCTCTTAACTATGATGATTTTCTTCCGATTTTATTA 720
 DB 661 ATTGTTTTTATATACATTTTCTTCTCTTAACTATGATGATTTTCTTCCGATTTTATTA 720
 QY 721 ATGACTATAAGTATTTTATATTAAGAGCAGCATGCGTATGTTGCGTCAAAAAATC 780
 DB 721 ATGACTATAAGTATTTTATATTAAGAGCAGCATGCGTATGTTGCGTCAAAAAATC 780
 QY 781 TTTCTGATTTTAAAGCTAGTTTGGCAACCCGTGTTCTTCAAGAAATTTGATTTT 840
 DB 781 TTTCTGATTTTAAAGCTAGTTTGGCAACCCGTGTTCTTCAAGAAATTTGATTTT 840
 QY 841 TTCAAAAAAATTAAGTATTTTCTTCTTAAATTAAGTATTTGATTTT 900
 DB 841 TTCAAAAAAATTAAGTATTTTCTTCTTAAATTAAGTATTTGATTTT 900
 QY 901 TGGCAGACTAGCCCTAGATGTTTCCCAATTAATTAATCAATCTGTGTAATTTATTTG 960
 DB 901 TGGCAGACTAGCCCTAGATGTTTCCCAATTAATTAATCAATCTGTGTAATTTATTTG 960
 QY 961 GCCAGCCCATTAATTTTAAACCGAACTGAATCGAGCAAAACCAATCTGAGCTAT 1020
 DB 961 GCCAGCCCATTAATTTTAAACCGAACTGAATCGAGCAAAACCAATCTGAGCTAT 1020
 QY 1021 TTCTCTAGATTTAAGAAAAAGAGAGAGAGAGAAATCAAGTTTAACTATTTGCC 1080
 DB 1021 TTCTCTAGATTTAAGAAAAAGAGAGAGAGAGAAATCAAGTTTAACTATTTGCC 1080
 QY 1081 TGAAGATGAGGTTTGGCAACGATAGCCGTAATATAGCTATAGGTCCTACGCTCA 1140
 DB 1081 TGAAGATGAGGTTTGGCAACGATAGCCGTAATATAGCTATAGGTCCTACGCTCA 1140
 QY 1141 GGTTCGAGCACTCTCGTGTATCTACATGAGCACTACATGCTTTGTTCAACGTTGCTC 1200
 DB 1141 GGTTCGAGCACTCTCGTGTATCTACATGAGCACTACATGCTTTGTTCAACGTTGCTC 1200
 QY 1201 TTGTTCCATGCTCGAAGCCTTTGCTTATTTGAACCAAGAGATCTACTCCAAACAT 1260
 DB 1201 TTGTTCCATGCTCGAAGCCTTTGCTTATTTGAACCAAGAGATCTACTCCAAACAT 1260
 QY 1261 CCATCTTACTCATGCAACTTCATGCAACGCAATATGTTTCTGTAAC 1311

DB 1261 CCATCTTACTCATGCAACTTCATGCAACGCAACGCAATATGTTTCTGTAAC 1311
 RESULT 5
 AAH76340
 ID AAH76340 standard; DNA; 255 BP.
 XX
 AC AAH76340;
 XX
 DT 29-OCT-2001 (first entry)
 XX
 DE Z. mays Ms45 promoter fragment.
 XX
 KW Ms45; male tissue; regulatory region; transcription; male fertility;
 KW hybrid seed; promoter; ds.
 XX
 OS Zea mays.
 XX
 PN W0200160997-A2.
 XX
 PD 23-AUG-2001.
 XX
 PF 13-FEB-2001; 2001WO-US004527.
 XX
 PR 15-FEB-2000; 2000US-00504487.
 XX
 PA (PION-) PIONEER HI-BRED INT INC.
 XX
 PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
 XX
 DR WPI; 2001-514772/56.
 XX
 PT A male tissue-preferred regulatory region comprising nucleotide sequences
 PT essential for initiating transcription of the Ms45 gene useful for
 PT mediating fertility in a male plant.
 XX
 PS Example 5; Fig 8; 50bp; English.
 XX
 CC The invention provides a male tissue-preferred regulatory region (I)
 CC comprising nucleotide sequences essential for initiating transcription of
 CC the Ms45 gene. A method of mediating male fertility in a plant is
 CC provided that involves introducing an expression vector comprising a
 CC promoter operably linked to (I) into a plant where the exogenous gene
 CC impacts male fertility of the plant and (I) controls expression of the
 CC exogenous gene. A method of producing hybrid seeds is also provided. The
 CC present sequence represents a Z. mays Ms45 promoter fragment
 XX
 SQ Sequence 255 BP; 59 A; 86 C; 39 G; 71 T; 0 U; 0 Other;
 Query Match 12.3%; Score 160.6; DB 5; Length 255;
 Best Local Similarity 97.6%; Pred. No. 1.3e-24;
 Matches 163; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1145 CGGCACTCTCGTGTATCTACATGAGCACTACATGCTGTTCAACGTTGCTGTG 1204
 DB 5 CGGCACTCTCGTGTATCTACATGAGCACTACATGCTGTTCAACGTTGCTGTG 64
 QY 1205 TCCATGTCGCAAGCCTTGGCTATTTGAAACCAAGAGATACCTACCCAAACATCCAT 1264
 DB 65 TCCATGTCGCAAGCCTTGGCTATTTGAAACCAAGAGATACCTACCTCCAAACATCCAT 124
 QY 1265 CTTACTCATGCAACTTCATGCAACGCAACATATGTTTCTGTAAC 1311
 DB 125 CTTACTCATGCAACTTCATGCAACGCAACATATGTTTCTGTAAC 171
 RESULT 6
 AAH76334
 ID AAH76334 standard; DNA; 158 BP.
 XX
 AC AAH76334;
 XX

DT 29-OCT-2001 (first entry)
XX
DE Z. mays Ms45 male tissue-preferred regulatory region fragment.
XX
KW Ms45; male tissue; regulatory region; transcription; male fertility;
KM hybrid seed; de.
XX
OS Zea mays.
XX
PN WO200160997-A2.
XX
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen WC, Fox TW, Garnat CW, Huffman G, Kendall TL;
XX WPI; 2001-514772/56.
XX
DR
XX
XX A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the Ms45 gene useful for
PT mediating fertility in a male plant.
XX
PS Claim 5; Page 47; 50pp; English.
XX
XX The invention provides a male tissue-preferred regulatory region (1)
CC comprising nucleotide sequences essential for initiating transcription of
CC the Ms45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (1) into a plant where the exogenous gene
CC impacts male fertility of the plant and (1) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a DNA fragment -38 to -195 bases upstream of
CC the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
CC nucleotide sequence
XX
SQ Sequence 158 BP; 41 A; 50 C; 21 G; 46 T; 0 U; 0 Other;
Query Match 11.1%; Score 146; DB 5; Length 158;
Best Local Similarity 99.4%; Pred. No. 1.5e-21;
Matches 157; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1155 CGTGTCACTCAGATGACATGATGCTGTTCAACGCTTCGTC-TTGTTCCATGTC 1213
DB 1 CGGTGATCTCAGATGACATGATGATGCTGTTCAACGCTTCGTTGTTCCATGTC 60
QY 1214 CAGACCTGCTTTCGAAAGAGATGATGCTGCTCCCAACATGCTTACTCAT 1273
DB 61 CAGACCTGCTTTCGAAAGAGATGATGCTGCTCCCAACATGCTTACTCAT 120
QY 1274 GCAACTTCCATGCAACAGCAGCATATGTTCTCTGAAC 1311
DB 121 GCAACTTCCATGCAACAGCAGCATATGTTCTCTGAAC 158
RESULT 7
ID AAL15210/c
XX AAL15210 standard; cDNA; 883 BP.
XX
AC AAL15210;
XX
DT 07-DEC-2001 (first entry)
XX
XX Human breast cancer expressed polynucleotide 7667.
DE Human breast cancer; cell marker; cyclostatic; ss.
KM Human; breast cancer; cell marker; cyclostatic; ss.
OS Homo sapiens.
XX

PN WO200151628-A2.
XX
XX 19-JUL-2001.
PD
XX
XX 10-JAN-2001; 2001WO-US000798.
PF
XX
XX 14-JAN-2000; 2000US-0176077P.
PR 14-MAR-2000; 2000US-0189167P.
PR 24-MAR-2000; 2000US-0192099P.
PR 29-MAR-2000; 2000US-0193480P.
PR 15-MAY-2000; 2000US-0205230P.
PR 09-JUN-2000; 2000US-0211315P.
PR 25-JUL-2000; 2000US-0220534P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Little J, Xu Y, Wang Y, Steinmann K;
XX
XX WPI; 2001-451856/48.
XX
XX
XX New peptide useful as a marker for the diagnosis of breast cancer.
PT
XX
XX Claim 1; Page 1378; 3695pp; English.
PS
XX
XX The invention relates to human breast cancer expressed polynucleotides
CC (AAL07544-AAL26789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterizing treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cyclostatic
CC activity
XX
SQ Sequence 883 BP; 322 A; 18 C; 23 G; 396 T; 0 U; 124 Other;
Query Match 5.4%; Score 70.2; DB 4; Length 883;
Best Local Similarity 39.2%; Pred. No. 3.4e-05;
Matches 237; Conservative 0; Mismatches 367; Indels 0; Gaps 0;
QY 419 TGTTCACATATGCGACCTACCCGAAACAGATTAAACAAAGTAATCAATCA 478
DB 873 TGTNNANNANACCTANACCTTTAAANATNTNNANNNTTAANNAAAAATATTAAANN 814
QY 479 CTCGAAGATCAATGATATGTTTAAAGAAACATCTATTAACACGATCCTTTAAA 538
DB 813 TAAATAATTTTAAATACTAATTTTAANNATTAATTTTNTTAAATAAAAAANNNAAAA 754
QY 539 AACAAGCATATTTGCAAGAGACAATTAATGTTACGTTTACAAACATCTAAGACGACA 598
DB 753 TTAANNNTTTTAAANTTAATTAACCAAAATTTTTTTTAAAAAATTTTTTAAAAANTT 694
QY 599 AATTATATCGAAAGGTAAAGCTATGACGTTACAGATTTTCTTTTCAATCTGTTATTGG 658
DB 693 AATAANTTAAATTTTAAATTAATTAATAAAAAAATTTTAAATTAATTAATTTTNTT 634
QY 658 TTAATGTTTAAATTAATTAATTTCTCTCTTAATTAATGATGTTTCTTCGATTTTAA 718
DB 633 TTTNTTNNATTAATAAANNTTTTAAATTAATAAANNTTTTNTTAAATTAATTAATTA 574
QY 719 AATGACTATAAAGTATTTTATTAATAAGACGACGATGCTGATGTTCTGTCAAAAA 778
DB 573 AAAAAAANNTTTTNTTTTAAANTTAATTAATAAATTTTATTTNTTTTCTTAAAAANA 514
QY 779 TCTTTCTGATTTTAAAGAGCTAGTTGGCAACCTGTTCTTTCAAGAAATTTGATT 838
DB 513 AAAAAAATTTTAAANNTTTTAAACCTTAATTAATAAANNTTTTNTTNTT 454
QY 839 TTTTCAAAAAAATAGTTTATTTCTCTTAAATTAATAAGAAACATTAAGAAAAATGA 898
DB 453 NNAANNTTAAATTTATTTTNTTTTAAATTTTAATTAATAAANNTTAATTAATTAAT 394
QY 899 GTTGCCAGACTAGCCCTAGAAATGTTTCCCAATTAATTAATCAATCACTGTGATTAATT 958

Db 393 TTTNAAAAATTAATTTATATATATATNNAANAATTTAAATNANANAAAAATTTTA 334
 QY 959 TGGCCAGCCCCATTAATTAATTTAAACCGAACTGAAATCGAGCAACCAATCTGAGCT 1018
 Db 333 NTTTAAATTAATTAATTAATTTAAACNTCATTTNTTAATTTTAATTAATTAATTAATTAAT 274
 QY 1019 ATTT 1022
 Db 273 NTTT 270

RESULT 8

ACN85231/C
 ID ACN85231 standard; DNA; 960 BP.
 XX ACN85231;
 XX 02-DEC-2004 (first entry)
 DE Breast cancer related marker, seq id 6381.
 XX
 XX Cancer; breast; tumour; cytostatic; marker; detection; therapy; dr.
 XX Homo sapiens.
 XX US2003099974-A1.
 XX 29-MAY-2003.
 XX 18-JUL-2002; 2002US-00198846.
 XX 18-JUL-2001; 2001US-0306220P.
 XX (MIL-) MILLENNIUM PHARM INC.
 XX Lillie J, Xu Y, Wang Y, Steinmann K;
 XX WPI; 2003-787014/74.
 XX
 PT Novel isolated polypeptide associated with breast cancer, useful for
 PT detecting presence of polypeptide in sample, as a marker for breast
 PT cancer.
 XX
 PS Disclosure, SEQ ID NO 6381; 36bp; English.

XX The invention relates to an isolated polypeptide (I) associated with
 CC breast cancer which is encoded by a nucleic acid molecule comprising a
 CC nucleotide sequence (S1). Further disclosed is an antibody that binds to
 CC the polypeptide of the invention. The activity of the polypeptide of the
 CC invention may be described as cytostatic. The antibody is useful for
 CC detecting the presence of (I) in a sample. Nucleic acid molecules of the
 CC invention are useful in the detection of breast tumours. (I) is useful as
 CC a marker for breast cancer and in breast cancer therapy. Sequences given
 CC in records ACN78851-ACN92934 represent nucleic acid markers associated
 CC with breast cancer. Note: The sequence listing does not form part of the
 CC specification but may be obtained in electronic format from the USPTO web
 CC site at seqdata.uspto.gov/sequence.html?docid=20030099974
 CC

SQ Sequence 960 BP; 340 A; 33 C; 39 G; 421 T; 0 U; 127 Other;

Query Match 5.4%; Score 70.2; DB 11; Length 960;
 Best Local Similarity 39.2%; Pred. No. 3.5e-05;
 Matches 237; Conservative 0; Mismatches 367; Indels 0; Gaps 0;

QY 419 TGTCACTATTGGCACTACCGAACAAGATTTAAAAAATTAACCAAGTACTAATCCA 478
 Db 933 TGTNNANNANNACTTAACCTTTAAANATTTNNANNTTNAANAAAAAATATTTAAAN 874
 QY 479 CTGAAAGCTATCATGTATGTTTAAAGAAACATCTATTAAACCAAGATCTTTAAA 538
 Db 873 TAAATAATTTTAAATAATTTTAAANATTAATTTTNTTAAATAATTTTAAATAATTTTAA 814

QY 539 AACAGCATATTTTCGAAGAAGACAAATTAATGTTACAGTTTACAAACATCTAAGAGACA 598
 Db 813 TTAANNNTTTTNTAANTTAATTAACCAAAATTTTTTTTTTAAAAAATTTTAAANNT 754
 QY 599 AATTATATCGAAAGTAGATGAGCTTACAGATTTTCTTTTCATCTGTGTTTGG 658
 Db 753 AATTAATTTAATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTT 694
 QY 659 TTAATGTTTAAATACATTTTCTTCTTACAAATAGATGATTTTCTCGATTTTATA 718
 Db 693 TTTNTTTNNAATAAAAAAATTTTAAATAAAAAAANNTTTTATNANATATTTAA 634
 QY 719 AATGACTATTAAGTATTTTATTAATGAAGCAGATGTCGAGATCTCGTCAAAA 778
 Db 633 AAAAAAANNTTTTTTTTAAATTAATTAATTAATTAATTAATTAATTTTCTNAAAA 574
 QY 779 TCTTTCGATTTTAAAGAGCTAGTTTGGCAACCTGTTTCTTCAAGAAATTTGATT 838
 Db 573 AAAAAAATTTTNTAANNTTTTAAACCTTNAATTTAANANAAATTTTNTNTT 514
 QY 839 TTTTCAAAAAAATTAATTTTCTTTTAAATTAATTAATTAATTAATTAATTAATTA 898
 Db 513 NNAATTTAATAAATTAATTTTNTTCTAATTAATTAATTAATTAATTAATTAATTAAT 454
 QY 899 GTTCCAGACTAGCCTAGAAATGTTTCCCAATTAATTAATCACTGCTATTAATT 958
 Db 453 TTTNAAAAAATATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 394
 QY 959 TGGCCAGCCCCATTAATTTAATTAACGAACGAAATCGAGGAAACCAATCTGAGCT 1018
 Db 393 NTTTAAATTAATTAATTAATTTAAACNTCATTTNTTAATTTTAATTTAAATAATTT 334
 QY 1019 ATTT 1022
 Db 333 NTTT 330

RESULT 9

ID ADR04296
 ID ADR04296 standard; DNA; 13400 BP.
 XX
 AC ADR04296;
 XX
 DT 04-NOV-2004 (first entry)
 XX
 XX Corn FT homologue nucleotide sequence SEQ ID NO:63.

XX
 XX flowering locus T gene; FT; terminal flower; TFL; Apetala3; Ap3; plant;
 XX floral development; plant sterility; plant fertility; flowering time;
 XX plant growth rate; inflorescence architecture; tissue culture morphology;
 XX cell division; FT homologue; gene; ds.
 XX
 OS Zea mays.
 XX
 PN MO2004067723-A2.
 XX
 PD 12-AUG-2004.
 XX
 PF 29-JAN-2004; 2004WO-US002422.
 XX
 PR 30-JUN-2003; 2003US-00343477.
 XX
 PA (PION-) PIONEER HI-BRED INT INC.
 PA (DUPO) DU PONT DE NEMOURS & CO E I.
 XX
 PI Danilevskaya O, Hermon P, Bruggemann E, Shitbroun D, Ananiev E;
 PI Rafalski JA, Sakai H, Cahoon E, Cahoon R, Klein T;
 XX
 DR WPI; 2004-580996/56.

XX New polynucleotides, specifically nucleic acid fragments encoding
 PT flowering locus T gene (FT) or terminal flower (TFL), or Apetala3 (Ap3)
 PT homologs, useful for floral development, e.g. engineering plant flowering

PT time.
PS Claim 6; SEQ ID NO 63; 109pp; English.
XX
XX The present invention describes an isolated polynucleotide comprising a
CC first, second, third, fourth or fifth nucleotide sequence, or their
CC complement encoding a polypeptide either having flowering locus T gene
CC (FT), terminal flower (FTL), or Apetala3 (Ap3) homologue locus T gene
CC described: (1) a vector comprising the polynucleotide; (2) a recombinant
CC DNA construct comprising the polynucleotide; (3) transforming a cell by
CC transforming a cell with the polynucleotide; (4) a cell comprising the
CC recombinant DNA construct; (5) producing a plant comprising transforming
CC a plant cell with the polynucleotide, and regenerating a plant from the
CC transformed plant cell; (6) a plant comprising the recombinant DNA
CC construct; (7) a seed comprising the recombinant DNA construct; (8) an
CC isolated polynucleotide comprising a first nucleotide sequence, where the
CC first nucleotide sequence contains at least 30 nucleotides, and where the
CC first nucleotide sequence is comprised by another polynucleotide, where
CC the other polynucleotide includes the second, third, fourth, fifth or
CC sixth nucleotide sequence; (9) an isolated polypeptide having FT or Ap3
CC homologue activity, as described above; and (10) isolating a polypeptide
CC encoded by the polynucleotide comprising isolating the polypeptide from a
CC cell containing a recombinant DNA construct comprising the polynucleotide
CC operably linked to a regulatory sequence. The polynucleotides are useful
CC for floral development, e.g. engineering plant sterility/fertility,
CC flowering time, plant growth rate, inflorescence architecture, and tissue
CC culture morphology and the rate of cell division to enhance
CC transformation. The present sequence represents an FT homologue
CC nucleotide sequence from the present invention.
XX
SQ Sequence 13400 BP; 3962 A; 2948 C; 2695 G; 3795 T; 0 U; 0 Other;
Query Match 5.3%; Score 69.8; DB 13; Length 13400;
Best Local Similarity 73.6%; Pred. No. 8e-05;
Matches 89; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
QY 811 ACCCTGTTCTTCAAGAAATTTGATTTTCMAAAAAAATAGTTTATTTCTCTTA 870
DB 8559 ACCATTATTTTCAAGAGGTTTTTATTTATCAAGAAAAATAGTCAATTTCTCTTGG 8618
QY 871 TAAATATGAAACACTTAGAAAAATAGATGGCCAGACTAGCCCTAGAAATGTTTCCAA 930
DB 8619 AAAAAATGAAATTCATTAGAAAAATGGGTTTCAAACTAGTCTTATTTAGTTTCCAT 8678
QY 931 T 931
DB 8679 T 8679
RESULT 10
ID AA210551
AA210551 standard; DNA; 2657 BP.
XX
AC AA210551;
XX
DT 16-NOV-1999 (first entry)
XX
DE DNA sequence of the P-Zeap promoter of maize.
XX
KW P-Zeap promoter; male sterile plant; glyphosate tolerance; glyphosate;
KW male reproductive tissue; hybrid seed production; crop outcrossing;
KW flower life; ss.
XX
OS Zea mays.
XX
PN WO946396-A2.
XX
PD 16-SEP-1999.
XX
PF 09-MAR-1999; 99WO-US005126.
XX
PR 09-MAR-1998; 98US-007277P.
XX

PA (MONS) MONSANTO CO.
XX
XX PI Brown SM, Fromm ME;
XX
XX DR WPI; 1999-551420/46.
XX
XX PT Production of male sterile plants using a gene encoding glyphosate
XX tolerance, used for, e.g. production of hybrid seed.
XX
XX PS Disclosure; Fig 1A-B; 54pp; English.
XX
XX The present sequence represents the P-Zeap promoter of maize. It is used
CC in the method of the invention. The specification describes a method for
CC the production of male sterile plants. The method comprises selective
CC expression of DNA encoding a protein that causes tolerance to glyphosate
CC and application of glyphosate. The method uses two DNA molecules, each
CC operably linked to a separate promoter, whereby the first promoter
CC functions in plant cells to produce a protein that causes tolerance to
CC glyphosate, and the second promoter functions in plant cells to cause the
CC production of a second RNA sequence in a male reproductive tissue.
CC Expression of the DNA promotes tolerance to glyphosate in those tissues
CC in which it is expressed. Expression of the second DNA molecule causes
CC the production of an RNA sequence which can inhibit the glyphosate
CC tolerance generated by expression of the first DNA molecule. By using a
CC promoter for the second DNA molecule which restricts the production of
CC the antisense RNA to only a subset of the tissues which express the first
CC DNA molecule, only the subset of tissues in which the second DNA molecule
CC is expressed will be susceptible to glyphosate toxicity. In this way, a
CC specific cell type or combination of cell types, depending upon the
CC promoters utilized, can be selectively ablated by application of
CC glyphosate to the plant. The methods can be used for producing male-
CC sterile plants for use in the production of hybrid seed, for minimizing
CC undesirable crop outcrossing, and for lengthening flower life. The
CC methods can be used with plants such as corn, wheat, rice, canola, oat,
CC barley, alfalfa, carrot, cotton, oilseed, oilseed rape, sugarbeet,
CC sunflower, soybean, tomato, cucumber and squash
XX
SQ Sequence 2657 BP; 686 A; 611 C; 677 G; 683 T; 0 U; 0 Other;
Query Match 5.1%; Score 67.2; DB 2; Length 2657;
Best Local Similarity 74.2%; Pred. No. 0.00019;
Matches 98; Conservative 0; Mismatches 33; Indels 1; Gaps 1;
QY 787 ATTTTAAAGAGCAGATTGGACCCGTTCTTCAAGAAATTTGATTTTCAAA 846
DB 1883 ATCTTTTAAAGGCTAGTTTGGAAACACATT-TTCAAGGATTTCAATTTTGCAG 1941
QY 847 AAAAATGATTTTCTCTTATGAAAAATGAAAAACACTTAGAAAAATAGATTGCCAG 906
DB 1942 GGAATTAAGTTTCATTTTCCCTTGGGAAATGAAATCCATGGAATAATGTGTTCCAA 2001
QY 907 ACTAGCCCTAGA 918
DB 2002 ACTAGCCCTAAA 2013
RESULT 11
ID AAX58751/C
AAX58751 standard; cDNA; 6027 BP.
XX
AC AAX58751;
XX
DT 16-AUG-1999 (first entry)
XX
DE Maize dulli gene encoding starch synthase enzyme DUL.
XX
KW Starch synthase; SSII; DUL; dulli gene; maize; transgenic plant; ss.
XX
OS Zea mays.
XX
XX Key Location/Qualifiers
FH misc_feature 1..1437
FT

```
FT /note= "functional fragment of cDNA specifically claimed
PT in Claim 16"
PT CDS 120..5147
PT /*cag= a
PT misc_feature 120..1221
PT /*cag= b
PT /note= "functional fragment of cDNA specifically claimed
PT in Claim 12"
FT misc_feature 565..816
FT /*cag= d
FT /note= "functional fragment of cDNA specifically claimed
FT in Claim 14"
FT misc_feature 655..1221
FT /*cag= c
FT /note= "functional fragment of cDNA specifically claimed
FT in Claim 13"
FT misc_feature 1369..1944
FT /*cag= e
FT /note= "functional fragment of cDNA specifically claimed
FT in Claim 15"
FT misc_feature 1438..2424
FT /*cag= g
FT /note= "functional fragment of cDNA specifically claimed
FT in Claim 17"
FT misc_feature 2425..3791
FT /*cag= h
FT /note= "functional fragment of cDNA specifically claimed
FT in Claim 18"
PT MO924575-A1.
PN 20-MAY-1999.
XX
XX 12-NOV-1998; 98WO-US024225.
XX
XX 12-NOV-1997; 97US-00968542.
XX
XX (IOWA ) UNIV IOWA STATE RES FOUND INC.
XX
XX Myers AM, James MG;
XX WPI; 1999-327406/27.
XX P-PSDB; AAY06199.
XX
XX Nucleic acid encoding starch synthase enzyme of maize.
XX
XX Claim 1; Page 104-107; 138pp; English.
XX
XX This is the nucleotide sequence of the maize gene dnl1 (dnl). To
XX illustrate the role of the dnl locus in starch biosynthesis, a transposon
XX tagging strategy was used to isolate the gene and describe its
XX polypeptide product. The invention reports tagging of the dnl locus with
XX C Mu transposon, cloning and characterization of a portion of the gene, and
XX isolation of a near full-length cDNA (the present sequence). The amino acid
XX sequence (see AAY06199) deduced from this cDNA indicates that Dnl codes
XX for a 186 kDa protein extremely similar to potato tuber starch synthase
XX SSII1. Its expression pattern indicates that Dnl codes for SSII of maize
XX endosperm. The Dnl product contains unique sequence features in its N-
XX terminus that may mediate direct interactions with other starch
XX CC biosynthetic enzymes. Mutations within the maize SSII gene affect
XX multiple aspects of starch biosynthesis by disrupting an enzyme complex
XX containing starch synthase(s), starch branching enzyme(s) and possibly
XX starch debranching enzyme(s). The isolated cDNA can be used to provide an
XX enzyme with which to regulate the production of starch, and with which to
XX produce altered or novel forms of starch, e.g. in transgenic plants.
XX Expression of Dnl in bacteria and yeasts also modifies glycogen
XX production. Claimed expression vectors comprise the cDNA or fragments of
XX it that code for functional portions of Dnl
XX
XX Sequence 6027 BP; 1800 A; 1177 C; 1441 G; 1609 T; 0 U; 0 Other;
```

```
Query Match 5.1%; Score 66.4; DB 2; Length 6027;
Best Local Similarity 77.5%; Pred. No. 0.00035;
```

```
Matches 93; Conservative 0; Mismatches 26; Indels 1; Gaps 1;
Qy 792 TTTAAGAGCTAGTTGGCAACCTGTTCTTTCAAGAAATTTGATTTTCAAAAAA 851
Db 5606 TCTACGGGCTAGTTGGAAACCCCATTT-TTCCAAGGATTTTCATTTTCCAAGAAAA 5548
Qy 852 TTAGTTATTTTCTCTTATTAATAATGAAAAACCTGGAATAATAGTTGCCAGCTAG 911
Db 5547 TTAGTTATTTTTCATTGAAAAAATTGAAATCTTTGAAAAAATAGATTCACTACTAG 5488
```

```
RESULT 12
ABX0935/C
ID ABX0935 standard; DNA; 6027 BP.
```

```
AC ABX0935;
XX
DT 17-FEB-2003 (first entry)
```

```
DE DNA encoding maize Starch synthase III (Dnl).
```

```
XX Starch; starch synthase; glucan association domain; GLASS; linker domain;
XX LINKR; glucosyl transferase domain; GLYTR; C-terminal end; CTEND;
XX granule bound starch synthase; GBS; morphology; retrogradation;
XX waterbinding; swelling potential; gene; db.
```

```
XX Zea mays.
```

```
XX MO200279410-A2.
```

```
XX 10-OCT-2002.
```

```
XX 29-MAR-2002; 2002WO-US009574.
```

```
XX 30-MAR-2001; 2001US-0279720P.
```

```
XX (BAD1 ) BASF PLANT SCI GMBH.
```

```
XX Commuri P, Keeling PL, Ramirez N, McKean A, Gao Z, Guan H;
```

```
XX WPI; 2003-040678/03.
```

```
PT New DNA encoding fusion protein consisting of 4 different functional
PT domains selected from glucan association domain, linker domain, glucosyl
PT transferase domain, and C-terminal end, useful for producing modified
PT starches.
```

```
PS Claim 33; Page 225-227; 265pp; English.
```

```
XX The invention describes an isolated DNA molecule encoding a fusion
XX protein consisting of 4 different functional domains selected from glucan
XX association domain (GLASS), linker domain (LINKR), glucosyl transferase
XX domain (GLYTR), and C-terminal end (CTEND) which are operably linked to
XX one another. The DNA molecule is useful for expressing in plants
XX polypeptides including starch synthase enzymes as fusion proteins with
XX improved affinity to starch and modified catalytic capabilities and to
XX CC the in vivo and in vitro synthesis of glucan chains of modified lengths
XX as compared to plants producing native starch or starch produced with
XX CC native starch synthases. Expression of the starch synthase fusion
XX CC protein along with granule bound starch synthase (GBSS) will lead to a
XX modified starch having an altered or improved morphology, retrogradation,
XX waterbinding, or swelling potential of the granules, gel strength,
XX CC adhesiveness, cohesiveness, hardness, elasticity, increased or decreased
XX granule size, degree of branching, crystallinity, degree of cross-
XX CC linking, and increased or decreased glucan chain lengths. This sequence
XX encodes a starch synthase used in the invention
```

```
XX Sequence 6027 BP; 1800 A; 1177 C; 1441 G; 1609 T; 0 U; 0 Other;
```

```
Query Match 5.1%; Score 66.4; DB 8; Length 6027;
Best Local Similarity 77.5%; Pred. No. 0.00035;
Matches 93; Conservative 0; Mismatches 26; Indels 1; Gaps 1;
```


QY 792 TTAAAGCTAGTTGGCAACCGTCTTTCAGAAAGATTTGATTTTTCAGAAAAA 851
DB 5606 TCTACGGGCTAGTTGGGAACCCCATTT-TCGACAGGATTTTCATTTTCCAGAAAA 5548
QY 852 TTAGTTATTTTCTCTTTATAAATAAGAAAAACATTAGAAAAATAGAGTTGCCAGACTAG 911
DB 5547 TTAGTTATTTTCTCTTTATAAATAAGAAAAACATTAGAAAAATAGAGTTGCCAGACTAG 5488

RESULT 13
ADK12106/c
ID ADK12106 standard; cDNA; 6027 BP.
XX
AC ADK12106;
XX
DT 03-JUN-2004 (first entry)
XX
DE cDNA encoding maize starch synthase, DUL.
XX
KM Maize; dul1 1; DUL; starch synthase;
KM alpha-1,4-glycosyltransferase catalytic activity; starch production;
KM plant; gene; ss.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT CDS 120..5144
FT /*tag= a
FT /product= "DUL"

XX US2004049810-A1.
XX
XX 11-MAR-2004.
XX
XX 05-AUG-2003; 2003US-00634262.
XX
XX 12-NOV-1997; 97US-00968467.
XX
XX 12-MAY-2000; 2000US-00544467.
XX
XX (MYER/) MYERS A M.
XX (JAME/) JAMES M G.
XX
XX Myers AM, James MG;
XX
XX MPI: 2004-238526/22.
XX
XX P-PSDB; ADK12117.
XX
XX New nucleic acid designated dul11, encoding a starch synthase, useful in
XX starch production.
XX
XX Claim 1; SEQ ID NO 1; 58pp; English.

XX The present invention relates to the isolation of a maize gene, dul1 1
XX (DUL), and the polypeptide it encodes. The DUL polypeptide has starch
XX synthase activity, and comprises an N-terminal arm region, a C-terminal
XX catalytic region, and a region of about 900 amino acids terminating with
XX the catalytic region. The C-terminal catalytic region has a catalytic
XX domain comprising alpha-1,4-glycosyltransferase catalytic activity. The
XX dul1 polynucleotide sequence is useful in producing starch e.g. from a
XX transgenic plant or transfected cell. The present sequence encodes maize
XX DUL.
XX
XX Sequence 6027 BP; 1800 A; 1177 C; 1441 G; 1609 T; 0 U; 0 Other;

QY Query Match 5.1%; Score 66.4; DB 12; Length 6027;
DB Best Local Similarity 77.5%; Pred. No. 0.00035;
Matches 93; Conservative 0; Mismatches 26; Indels 1; Gaps 1;
QY 792 TTAAAGCTAGTTGGCAACCGTCTTTCAGAAAGATTTGATTTTTCAGAAAAA 851
DB 5606 TCTACGGGCTAGTTGGGAACCCCATTT-TCGACAGGATTTTCATTTTCCAGAAAA 5548
QY 852 TTAGTTATTTTCTCTTTATAAATAAGAAAAACATTAGAAAAATAGAGTTGCCAGACTAG 911

DB 5547 TTAGTTATTTTCTCTTTATAAATAAGAAAAACATTAGAAAAATAGAGTTGCCAGACTAG 5488

RESULT 14
ABX35844
ID ABX35844 standard; cDNA; 439 BP.
XX
XX ABX35844;
XX
XX 20-FEB-2003 (first entry)
XX
DE Bovine EST associated with lactation/muscle/fat deposition #1009.
XX
XX Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
XX muscle deposition; fat deposition; genome mapping; gene identification;
XX gene analysis; cattle breeding.
XX
OS Bos Taurus.
XX
XX US200237139-A1.
XX
XX 26-SEP-2002.
XX
XX 24-SEP-2001; 2001US-00960352.
XX
XX 12-JAN-1999; 99US-0115707P.
XX
XX 11-JAN-2000; 2000US-00480902.
XX
XX (BYAT/) BYATT J C.
XX (MATH/) MATHIALAGAN N.
XX (TAON/) TAO N.
XX (WAR/) WARREN W C.
XX
XX Byatt JC, Mathialagan N, Tao N, Warren WC;
XX
XX MPI: 2003-110599/10.
XX
XX New nucleic acid associated with lactation, and muscle and fat
XX deposition, useful for genome mapping, gene identification and analysis,
XX cattle breeding, or for genetically improving cattle.
XX
XX Claim 2; SEQ ID NO 1009; 245pp; English.

XX The invention relates to a purified nucleic acid molecule associated with
XX lactation or muscle and fat deposition (designated LMFD), derived from
XX cattle, and the LMFD nucleic acid can specifically hybridise to a second
XX nucleic acid molecule comprising any of 15112 nucleotide sequences,
XX appearing as ABX34836-ABX49947, or complements of them. Also included are
XX : (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
XX acid linked to a promoter and a 3' non- translated sequence that
XX functions in the cell to cause termination of transcription and addition
XX of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
XX (2) determining a level or pattern of a molecule in a bovine cell or
XX tissue comprising: (a) incubating a marker nucleic acid (comprising any
XX of the 15112 nucleic acid sequences or its complement or fragment) with a
XX complementary nucleic acid molecule obtained from the bovine cell or
XX tissue, where hybridisation between the marker nucleic acid and the
XX complementary nucleic acid permits the detection of the molecule; and (b)
XX detecting the level or pattern of the complementary nucleic acid, where
XX the detection of the complementary nucleic acid is predictive of the
XX level or pattern of the molecule. The LMFD nucleic acid is used for
XX determining a level or pattern of a molecule in a bovine cell or tissue.
XX It is useful for genome mapping, gene identification and analysis, cattle
XX breeding, preparation of constructs for use in cattle gene expression, or
XX for genetically improving cattle. The present sequence is one of the
XX 15112 bovine LMFD ESN (expressed sequence tag) nucleic acids. Note: The
XX present sequence was not shown in the specification but was obtained in
XX electronic format from the USPTO web site:
XX seqdata.uspto.gov/sequence.html?docID=20020137139
XX
XX Sequence 439 BP; 45 A; 51 C; 56 G; 286 T; 0 U; 1 Other;

Query Match 5.0%; Score 66.2; DB 8; Length 439;
Best Local Similarity 54.3%; Pred. No. 0.0002;
Matches 134; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

QY 626 TTCAGATTTTCTTTTCATCTGTATTTGTATGTTTATATACATTTCTCT 685
DB 182 TTCGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCTT 241
QY 686 CTTACATAGAGTATTTCTCCGATTTTAAAGCATTAAGCTATTTATATA 745
DB 242 TTTTATT 301
QY 746 AGACAGCAGATGCTGATCTGTCAGAAATCTTCTGATTTTAAAGCTAGT 805
DB 302 TGTCGGGGGCTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 361
QY 806 TGGCAACCTGTTTCTTCAAGATTTGATTTTCAAAAAAATAGTTATTTCT 865
DB 362 TTTTTTTTTTTTTTTTTTTTTTTTTTTATTTATATCTTATTTTTTTTTT 421
QY 866 CTTTATA 872
DB 422 TTTCTTATA 428

RESULT 15
AA187279/C
ID AA187279 standard; cDNA; 346 BP.
XX
AC AA187279;
XX
XX 06-NOV-2001 (first entry)
DE Human polynucleotide SEQ ID NO 7339.
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KM vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KM tissue growth factor; immunomodulatory; cancer; leukemia;
KM nervous system disorders; arthritis; inflammation; se.
XX
OS Homo sapiens.
XX
XX WO200164835-A2.
PN
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US004927.
XX
PR 28-FEB-2000; 2000US-00515126.
PR 18-MAY-2000; 2000US-00577409.
XX
PA (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
PI WPI, 2001-514838/56.
DR P-PSDB; AAO07348.
XX
XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT and treating e.g. leukemia, inflammation and immune disorders.
XX
PS Claim 1; SEQ ID NO 7339; 1399bp + Sequence Listing; English.

CC The invention relates to human polynucleotides (AA179941-AA193841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukemia, nervous system disorders, arthritis and

CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX

SQ Sequence 346 BP; 188 A; 22 C; 16 G; 120 T; 0 U; 0 Other;

Query Match 5.0%; Score 65.6; DB 4; Length 346;
Best Local Similarity 52.6%; Pred. No. 0.00026;
Matches 143; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

QY 446 AGATTTAAAAAATTAACCAAGTACATATCCAGTGAAGCTATCATGTAATGTTTAA 505
DB 334 ACAATTAAGTATATATATTAATAATTAAGATTAATATTTATCATTAATAATTTTA 275
QY 506 GAAACATCTTATTAACCAAGATCTTTAAAAAACAAGCATATTTGAAAGACAAAT 565
DB 274 GACTATATATATATTAATAATTTTTTTTTTGAATAATTAATTTACATTTAAACAATAAAA 215
QY 566 TATGTTACGTTTACCAACCTTAAGAGCACAATATATCGAAAGGTAACTATGACG 625
DB 214 TTTTAAATATTAATAATTAATAATTAATTAATTAATAATAATTAATTTATTT 155
QY 626 TTCAGATTTTCTTTTCATCTGTATTTGTATGTTTATATATACATTTCTCT 685
DB 154 TATTTATT 95
QY 686 CTTACATAGAGTATTTCTCCGATTTTAT 717
DB 94 TTTTTTTTTTTTTTTTTTTTTTTTTTCTTTTTTT 63

Search completed: September 15, 2005, 21:33:05
Job time : 636.708 secs

This Page Blank (uspro)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 14, 2005, 19:14:19 ; Search time 198.304 Seconds
(without alignments)
10817.505 Million cell updates/sec

Title: US-10-713-381-1_COPY_1_1311

Sequence: 1 cccctggtcgtctcctcgaaga.....cgacacatcgtctcctcgaac 1311

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptcdatc/1/ina/5A COMB.seq:*
- 2: /cgn2_6/ptcdatc/1/ina/5B COMB.seq:*
- 3: /cgn2_6/ptcdatc/1/ina/6A COMB.seq:*
- 4: /cgn2_6/ptcdatc/1/ina/6B COMB.seq:*
- 5: /cgn2_6/ptcdatc/1/ina/PTCUS COMB.seq:*
- 6: /cgn2_6/ptcdatc/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1311	100.0	1394	3 US-08-880-499-1	Sequence 1, Appl1
2	1311	100.0	1394	3 US-08-880-499-2	Sequence 2, Appl1
3	66.4	5.1	6027	2 US-08-968-542C-1	Sequence 1, Appl1
4	66.4	5.1	6027	4 US-09-554-467A-1	Sequence 1, Appl1
5	58.6	4.5	2523	2 US-08-410-784A-3	Sequence 13, Appl1
6	57	4.3	19124	2 US-08-487-826B-13	Sequence 13, Appl1
7	55.8	4.2	279	4 US-09-313-294A-5397	Sequence 5397, Ap
8	55.6	4.2	612	4 US-09-902-540-1357	Sequence 1357, Ap
9	55.6	4.2	1394	3 US-08-880-499-1	Sequence 1, Appl1
10	55.6	4.2	1394	3 US-08-880-499-2	Sequence 2, Appl1
11	55.4	4.2	1055	4 US-09-806-708B-23	Sequence 23, Appl1
12	55.2	4.2	2614	4 US-09-004-056-1	Sequence 1, Appl1
13	55	4.2	396	4 US-09-640-173-53	Sequence 53, Appl1
14	55	4.2	396	4 US-09-713-550-53	Sequence 53, Appl1
15	55	4.2	396	4 US-09-825-294-53	Sequence 53, Appl1
16	55	4.2	396	4 US-09-970-866-53	Sequence 53, Appl1
17	54.6	4.2	55886	4 US-09-949-016-15129	Sequence 15129, A
18	53.8	4.1	307	4 US-09-313-294A-4743	Sequence 4743, Ap
19	53.8	4.1	6027	2 US-08-968-542C-1	Sequence 1, Appl1
20	53.8	4.1	6027	4 US-09-554-467A-1	Sequence 1, Appl1
21	52.2	4.0	1141	4 US-09-806-708B-22	Sequence 22, Appl1
22	51.6	3.9	134987	4 US-09-949-016-15348	Sequence 15348, A
23	51.6	3.9	134987	4 US-09-949-016-15349	Sequence 15349, A
24	51.6	3.9	134987	4 US-09-949-016-15350	Sequence 15350, A
25	51.6	3.9	134987	4 US-09-949-016-15507	Sequence 15507, A
26	51.6	3.9	134987	4 US-09-949-016-15508	Sequence 15508, A
27	51.6	3.9	134987	4 US-09-949-016-15509	Sequence 15509, A

ALIGNMENTS

C	28	51.4	3.9	50383	4	US-09-949-016-17600	Sequence 17600, A
C	29	51.4	3.9	129415	4	US-09-949-016-16997	Sequence 16997, A
C	30	51	3.9	601	4	US-09-949-016-156535	Sequence 156535, A
C	31	51	3.9	1141	4	US-09-806-708B-22	Sequence 22, Appl1
C	32	50.6	3.9	1039	4	US-09-902-540-1280	Sequence 1280, Ap
C	33	50.6	3.9	16573	4	US-09-949-016-14876	Sequence 14876, A
C	34	50.6	3.9	18773	4	US-09-949-016-14164	Sequence 14164, A
C	35	50.4	3.8	731	1	US-08-451-405A-2	Sequence 2, Appl1
C	36	50.2	3.8	832	4	US-09-621-976-2813	Sequence 2813, Ap
C	37	50.2	3.8	317366	4	US-09-949-016-16001	Sequence 16001, A
C	38	50	3.8	1392	3	US-09-257-584-1	Sequence 1, Appl1
C	39	49.8	3.8	2435	4	US-09-306-593-1	Sequence 1, Appl1
C	40	49.8	3.8	231129	4	US-09-949-016-16110	Sequence 16110, A
C	41	49.8	3.8	266283	4	US-09-949-016-11934	Sequence 11934, A
C	42	49.6	3.8	640681	4	US-09-790-988-1	Sequence 1, Appl1
C	43	49.2	3.8	126176	4	US-09-949-016-16137	Sequence 16137, A
C	44	49.2	3.8	126176	4	US-09-949-016-16138	Sequence 16138, A
C	45	47.8	3.6	601	4	US-09-949-016-25787	Sequence 25787, A

RESULT 1
US-08-880-499-1
Sequence 1, Application US/0880499

Patent No. 6037523

GENERAL INFORMATION:

APPLICANT: Alberson, Marc C.

APPLICANT: Fox, Tim W.

APPLICANT: Carl, Garnat W.

APPLICANT: Huffman, Gary A.

APPLICANT: Kendall, Timmy L.

TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION

TITLE OF INVENTION: AND METHOD OF USING SAME

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.

STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.

STREET: Box 1000

CITY: Johnston

STATE: Iowa

COUNTRY: USA

ZIP: 50131

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/880,499

FILING DATE: CONCURRENTLY HEREWITH

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Sweeney, Patricia A.

REGISTRATION NUMBER: 32,733

REFERENCE/DOCKET NUMBER: 0578

TELECOMMUNICATION INFORMATION:

TELEPHONE: (515) 248-4800

TELEFAX: (515) 248-4844

INFORMATION FOR SEQ. ID NO. 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1394 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-880-499-1

Query Match 100.0%; Score 1311; DB 3; Length 1394;

Best Local Similarity 100.0%; Pred. No. 1,1e-313;

Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAATGCTCTCTATGAAAAAGATGATCAATGCTCTATATCCGTTTCTTAAGGCTC 60
DB 1 CCAATGCTCTCTATGAAAAAGATGATCAATGCTCTATATCCGTTTCTTAAGGCTC 60
QY 61 CTTCTTCTGCTTATTAAGTCACTGAAATCGGGGTTTACAAAAAATTCCAGGGTGATGAT 120
DB 61 CTTCTTCTGCTTATTAAGTCACTGAAATCGGGGTTTACAAAAAATTCCAGGGTGATGAT 120
QY 121 CTCATGCTCTCTCTGACCTCGGCTGACATTTCTGGAATGTCGGTGGTTCCTAT 180
DB 121 CTCATGCTCTCTCTGACCTCGGCTGACATTTCTGGAATGTCGGTGGTTCCTAT 180
QY 181 CTGACCGAGGCCCATGACACCTTTCCGAGCACTCCATCAAGGGCTTTTGGATGGCCCA 240
DB 181 CTGACCGAGGCCCATGACACCTTTCCGAGCACTCCATCAAGGGCTTTTGGATGGCCCA 240
QY 241 CGAAGCTATCGGGTCTGCTGATCCAGGGGATATATGTCGCCCAATGTCACCTATA 300
DB 241 CGAAGCTATCGGGTCTGCTGATCCAGGGGATATATGTCGCCCAATGTCACCTATA 300
QY 301 TTAATATCTTTAGATATTAATTAATTTTGAATAAATAAATACTTATCTTTGTTGTA 360
DB 301 TTAATATCTTTAGATATTAATTAATTTTGAATAAATAAATACTTATCTTTGTTGTA 360
QY 361 GGGCTCTGAGCATGATTTTCTTGGGCGCCAGAAATGCGAGGACGAGCATGCTTATG 420
DB 361 GGGCTCTGAGCATGATTTTCTTGGGCGCCAGAAATGCGAGGACGAGCATGCTTATG 420
QY 421 TCCACTTTTGGCACTACCCGAAACAGATTTTAAAAATAACCAAGTAACTATCCACT 480
DB 421 TCCACTTTTGGCACTACCCGAAACAGATTTTAAAAATAACCAAGTAACTATCCACT 480
QY 481 CGAAGCTATCATGTAATGTTTAAAGAAACATCTATTAACCAAGTAACTATCCACT 540
DB 481 CGAAGCTATCATGTAATGTTTAAAGAAACATCTATTAACCAAGTAACTATCCACT 540
QY 541 CAAGCATATTTGAAAGAGCAAAATTAATGTTACAGTTTACAAACATCTTAAGAGCAAA 600
DB 541 CAAGCATATTTGAAAGAGCAAAATTAATGTTACAGTTTACAAACATCTTAAGAGCAAA 600
QY 601 TTTATTCGAAAGGTAAGTATGACGTTCAATTTTCTTTTCAATCTTGTATTTTGT 660
DB 601 TTTATTCGAAAGGTAAGTATGACGTTCAATTTTCTTTTCAATCTTGTATTTTGT 660
QY 661 ATGTTTATATACATTTTCTCTCTTCAATAGATGATTTCTTCCGATTTTAA 720
DB 661 ATGTTTATATACATTTTCTCTCTTCAATAGATGATTTCTTCCGATTTTAA 720
QY 721 ATGACTATAAAGTCAATTTTATAAAGAGCAGCATGCTGATGATCTCGTCAAAAATC 780
DB 721 ATGACTATAAAGTCAATTTTATAAAGAGCAGCATGCTGATGATCTCGTCAAAAATC 780
QY 781 TTTCTGATTTTAAAGAGTATGTTGGCAACCTGTTCTTCAAGAAATTTGATTTT 840
DB 781 TTTCTGATTTTAAAGAGTATGTTGGCAACCTGTTCTTCAAGAAATTTGATTTT 840
QY 841 TTTCAAAAAAATAGTTATTTTCTCTTAAATAAGAAACATTAAGAAATAGAGT 900
DB 841 TTTCAAAAAAATAGTTATTTTCTCTTAAATAAGAAACATTAAGAAATAGAGT 900
QY 901 TGGCAGACTAGCCCTAGAAATGTTTCCCAATAAATTAATCAATCACTGATTAATTTG 960
DB 901 TGGCAGACTAGCCCTAGAAATGTTTCCCAATAAATTAATCAATCACTGATTAATTTG 960
QY 961 GGCAGGCCCAATTAATTTTAAACCGAAATCGAAATCGAGCGAAACCAATCTGAGCTAT 1020
DB 961 GGCAGGCCCAATTAATTTTAAACCGAAATCGAAATCGAGCGAAACCAATCTGAGCTAT 1020
QY 1021 TTTCTAGATTAAGTAAAAAGAGAGAGAGAGAGAAATCACTTTTAAGCTATGTTCC 1080
DB 1021 TTTCTAGATTAAGTAAAAAGAGAGAGAGAGAGAAATCACTTTTAAGCTATGTTCC 1080
QY 1081 TGAGATGTGGGTTTGGCAACGATAGCAACCGTAAATCATAGTCAATAGTGCCTAGCTCA 1140

DB 1081 TGAGATGTGGGTTTGGCAACGATAGCAACCGTAAATCAATAGCTCAATAGTGCCTAGCTCA 1140
QY 1141 GGTTCGGAGCTCTGCTGATCTCATCTCATGAGCATACTACTGTTTCAACCGTTGCTC 1200
DB 1141 GGTTCGGAGCTCTGCTGATCTCATCTCATGAGCATACTACTGTTTCAACCGTTGCTC 1200
QY 1201 TTTGTCATGCTCAAGCCCTGCTATTTCTGAACCAAGAGGATACCTACTCCCAACAT 1260
DB 1201 TTTGTCATGCTCAAGCCCTGCTATTTCTGAACCAAGAGATACCTACTCCCAACAT 1260
QY 1261 CCATCTTACTCATGCAACTTCCATGCAACACGACATATGTTTCTGAAAC 1311
DB 1261 CCATCTTACTCATGCAACTTCCATGCAACACGACATATGTTTCTGAAAC 1311

RESULT 2
US-08-880-499-2
; Sequence 2, Application US/08880499
; Patent No. 6037523
; GENERAL INFORMATION:
; APPLICANT: Albertson, Marc C.
; APPLICANT: Fox, Tim W.
; APPLICANT: Carl, Garmaat W.
; APPLICANT: Huffman, Gary A.
; APPLICANT: Kendall, Timmy L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION
; TITLE OF INVENTION: AND METHOD OF USING SAME
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
; STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.
; CITY: Johnston
; STATE: Iowa
; COUNTRY: USA
; ZIP: 50131
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/880,499
; FILING DATE: CONCURRENTLY HEREWITH
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Sweeney, Patricia A.
; REGISTRATION NUMBER: 32,733
; REFERENCE/DOCKET NUMBER: 0578
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (515) 248-4800
; TELEFAX: (515) 248-4844
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1394 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULAR TYPE: DNA (genomic)
US-08-880-499-2

Query Match 100.0%; Score 1311; DB 3; Length 1394;
Best Local Similarity 100.0%; Pred. No. 1.1e-313; Indels 0; Gaps 0;
Matches 1311; Conservative 0; Mismatches 0;

QY 1 CCAATGCTCTCTATGAAAAAGATGATCAATGCTCTATATCCGTTTCTTAAGGCTC 60
DB 1 CCAATGCTCTCTATGAAAAAGATGATCAATGCTCTATATCCGTTTCTTAAGGCTC 60
QY 61 CTTCTTCTGCTTATTAAGTCACTGAAATCGGGGTTTACAAAAAATTCCAGGGTGATGAT 120
DB 61 CTTCTTCTGCTTATTAAGTCACTGAAATCGGGGTTTACAAAAAATTCCAGGGTGATGAT 120

QY 121 CTCGATGTTCCACTTCTCCACCTGGGTTGCACTTCTTGATGTCGGTGTTCAT 180
DB 121 CTCGATGTTCCACTTCTCCACCTGGGTTGCACTTCTTGATGTCGGTGTTCAT 180
QY 181 CTGACCGAGGCCCATCAGACACTTCTGGGACACCCATCAAGGGCTTCTGSAATGCCA 240
DB 181 CTGACCGAGGCCCATCAGACACTTCTGGGACACCCATCAAGGGCTTCTGSAATGCCA 240
QY 241 CGAGACGATCGGGTCTGTGATCGAGGGGATATATGTCGCCCAATGTCACCTATA 300
DB 241 CGAGACGATCGGGTCTGTGATCGAGGGGATATATGTCGCCCAATGTCACCTATA 300
QY 301 TTTATTTCTTTAGATATTTATTTTAAATTTTGGAAAAATTAACAACTTATCTTTGCTA 360
DB 301 TTTATTTCTTTAGATATTTATTTTAAATTTTGGAAAAATTAACAACTTATCTTTGCTA 360
QY 361 GGGCTCAGCATAGATTTTGGCTTGGGGCCAGAAATGCGAGGACGACGATGCTAGTG 420
DB 361 GGGCTCAGCATAGATTTTGGCTTGGGGCCAGAAATGCGAGGACGACGATGCTAGTG 420
QY 421 TCCACTATTTGGCACTACCCAGAACAGATTTAAAAAATAACCAAGTAACTATCCACT 480
DB 421 TCCACTATTTGGCACTACCCAGAACAGATTTAAAAAATAACCAAGTAACTATCCACT 480
QY 481 CGAAAGCTATCATGTATGTTTAAAGAAACATCTATTTAAACCAAGTAACTATCCACT 540
DB 481 CGAAAGCTATCATGTATGTTTAAAGAAACATCTATTTAAACCAAGTAACTATCCACT 540
QY 541 CAAGCATATTTGCAAGAGACAAATTTATTTACAGTTTACAAACATCTAAGAGGACAA 600
DB 541 CAAGCATATTTGCAAGAGACAAATTTATTTACAGTTTACAAACATCTAAGAGGACAA 600
QY 601 TTAATGAAAGGTAGACATGACGTTTCAAGTTTCTTTTCAATTTCTGTATTTGTT 660
DB 601 TTAATGAAAGGTAGACATGACGTTTCAAGTTTCTTTTCAATTTCTGTATTTGTT 660
QY 661 ATTGTTTTTATATGATTTTCTCTCTTAAATGAGATTTTCTTCCGATTTTATATA 720
DB 661 ATTGTTTTTATATGATTTTCTCTCTTAAATGAGATTTTCTTCCGATTTTATATA 720
QY 721 ATGACATATAAGTCAATTTTATATTAAGAGACGATGCTGATGATCTGCTCAAAAATC 780
DB 721 ATGACATATAAGTCAATTTTATATTAAGAGACGATGCTGATGATCTGCTCAAAAATC 780
QY 781 TTTCTGATTTTTTAAAGCTAGTTTGGCAACCTGTTTCTTCAAGAAATTTGATTTT 840
DB 781 TTTCTGATTTTTTAAAGCTAGTTTGGCAACCTGTTTCTTCAAGAAATTTGATTTT 840
QY 841 TTTCAAAAAAATTAAGTTATTTCTCTTAAATGAAACCACTTAAAGAAATGAGT 900
DB 841 TTTCAAAAAAATTAAGTTATTTCTCTTAAATGAAACCACTTAAAGAAATGAGT 900
QY 901 TGGCAGACTAGCCCTAGATGTTTCCCAATTAATCACTGATGTAATTAATTTG 960
DB 901 TGGCAGACTAGCCCTAGATGTTTCCCAATTAATCACTGATGTAATTAATTTG 960
QY 961 GCGAGCCCATTAATTTTAAACCGAACTGAAATGAGGAAACCAATCTGAGCTAT 1020
DB 961 GCGAGCCCATTAATTTTAAACCGAACTGAAATGAGGAAACCAATCTGAGCTAT 1020
QY 1021 TTTCTAGATTTAGTAAAAAGGAGAGAGAGAAATCAAGTTTAACTATTTGCC 1080
DB 1021 TTTCTAGATTTAGTAAAAAGGAGAGAGAGAAATCAAGTTTAACTATTTGCC 1080
QY 1081 TGAGATGTCGGTGTGCAAGATGCGATCATGATGATGATGATGATGATGATGATGAT 1140
DB 1081 TGAGATGTCGGTGTGCAAGATGCGATCATGATGATGATGATGATGATGATGATGAT 1140
QY 1141 GGTTCGGACGCTCTGTCATCTGATGATGATGATGATGATGATGATGATGATGAT 1200
DB 1141 GGTTCGGACGCTCTGTCATCTGATGATGATGATGATGATGATGATGATGATGAT 1200

QY 1201 TTGTCATGTCGCAAGCCTTGCTATTCGAAACCAAGGATTAAGTACCTACCCAAACAT 1260
DB 1201 TTGTCATGTCGCAAGCCTTGCTATTCGAAACCAAGGATTAAGTACCTACCCAAACAT 1260
QY 1261 CCATCTTACATGCAACTTCCATGCAACACGACATATGTTTCTGTAAC 1311
DB 1261 CCATCTTACATGCAACTTCCATGCAACACGACATATGTTTCTGTAAC 1311

RESULT 3
US-08-968-542C-1/C
; Sequence 1, Application US/08968542C
; Patent No. 5981728
; GENERAL INFORMATION:
; APPLICANT: Myers, et al.
; TITLE OF INVENTION: dutil Codes For A No. 5981728e1 Search
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McGregor & Adler, LLP
; STREET: 8011 Candle Lane
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 floppy disk
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word 6.0.1 for Macintosh
; CURRENT APPLICATION DATA:
; FILING DATE: US/08/968,542C
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benjamin Aaron Adler, Ph.D., J.D.
; REGISTRATION NUMBER: 35,423
; REFERENCE/DOCKET NUMBER: D6036
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713) 777-2321
; TELEFAX: (713) 777-6908
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6027 bp
; TYPE: nucleic acid
; STRANDEDNESS: double-stranded
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: cDNA to mRNA
; HYPOTHETICAL: no
; ANTI-SENSE: no
; ORIGINAL SOURCE:
; ORGANISM: maize
; TISSUE TYPE: endosperm
; IMMEDIATE SOURCE:
; LIBRARY: (Gf11
; CLONE: pMg110; pMg6Aa; pMg6-2M
; US-08-968-542C-1

Query Match 5.1%; Score 66.4; DB 2; Length 6027;
Best Local Similarity 77.5%; Pred. No. 2.4e-06;
Matches 93; Conservative 0; Mismatches 26; Indels 1; Gaps 1;

QY 792 TTTAAGACTAGTTTGGCAACCTGTTCTTTCAAGAATTTTGAATTTTCAAAAAAA 851
DB 5606 TCTACGGGCTAGTTTGGGAACCCCAATTT-TTCCAAGGATTTTCATTTTCAAGAAAA 5548
QY 852 TTAGTTATTTTCTCTTTTAAATGAAAAACATTTGAAAAATGAGTTGCGAGACTAG 911

TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19124 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; US-08-487-826B-13

Query Match 4.3%; Score 57; DB 2; Length 19124;
Best Local Similarity 50.2%; Pred. No. 0.00078;
Matches 141; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY 614 TAAAGTATGACGTTGACATTTTCTTTTTCATCTGTATTTGTTATTTTATAT 673
DB 15677 TAAAGTATTTTCTTTCTTTGTTATTTTATATTTTATTTTATAT 15618
QY 674 ACATTTTCTCTTACATATGAGTATTTTCCGATTTTATAAATGACTATAAAGT 733
DB 15617 AAATTTTATTTTAAATTTTATTTTATATCTTTTCAATTTTATCTATCAAAATTTA 15558
QY 734 CATTTTATATAGAGACGACGATCTGATATCTGTCATAAAATCTTCTGATTTT 793
DB 15557 TATTTTATATATTTTATTTTATTTTAAAAAATTTTCTCTTTTATTTT 15498
QY 794 TAAAGCTAGTTGGCAACCTGTTCTTCAAAAGATTTTGAATTTTCAAAAAAAT 853
DB 15497 TTTTATTTTAAATTTTATTTTATTTTATTTTCAATTTTCTTTTTCATTTTAAATTT 15438
QY 854 AGTTTATTTCTCTTTTATAAATAGAAAACACTTAGAAAA 894
DB 15437 GTTTTATATTTCTTTTATAATTAATACATATATATAA 15397

RESULT 7
US-09-313-294A-5397
; Sequence 5397, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Laligudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 5397
; LENGTH: 279
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: incycle ID No. 6476212 700350078H1
; NAME/KEY: unsure
; LOCATION: 10, 12, 185-186, 204, 253, 274, 278
; OTHER INFORMATION: a, t, c, g, or other
; US-09-313-294A-5397

Query Match 4.3%; Score 55.8; DB 4; Length 279;
Best Local Similarity 69.0%; Pred. No. 0.00031;
Matches 89; Conservative 0; Mismatches 39; Indels 1; Gaps 1;
QY 789 TTTTAAAGCTAGTTGGCAACCTGTTCTTCAAAAGATTTGATTTTCAAAA 848
DB 3 TTCTTAANGNCCTAGTTGGAAACCCATTTTCCCGCGGTTTTCATTTTCCCAAGG 62
QY 849 AAATTAAGTTATTTCTCTTAT-AAAATGAAAAAAGCTAGAAAAATAGTTGCAGA 907

DB 63 AAGTAGAACATTTTCCCTTGGAATAATAGAAATCCTTGGAAAAATCGATTTCCCAA 122
QY 908 CTAGCCCTA 916
DB 123 CTAGCCCTA 131

RESULT 8
US-09-902-540-1357/c
; Sequence 1357, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Miegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1357
; LENGTH: 612
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(612)
; OTHER INFORMATION: unsure at all n locations
; US-09-902-540-1357

Query Match 4.2%; Score 55.6; DB 4; Length 612;
Best Local Similarity 51.4%; Pred. No. 0.00047;
Matches 127; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY 626 TTCAGATTTTCTTTTCAATCTGTTATTTGTTATGTTTATATACATTTCTTCT 685
DB 560 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 501
QY 686 CTAGCAATGAGATTTCTCCGATTTTATTAATAGCTATTAAGCTATTTATATA 745
DB 500 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 441
QY 746 AGAGCAGCATGTCGATTCCTGTCAAAAATCTTCGATTTTAAAGCTAGTT 805
DB 440 ATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 381
QY 806 TGGCAACCTGTTCTTCAAAAGATTTGATTTTCAAAAAAATAGTTATTTCT 865
DB 380 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 321
QY 866 CTTTATA 872
DB 320 TATTTTA 314

RESULT 9
US-08-880-499-1/c
; Sequence 1, Application US/08880499
; Patent No. 6037523
; GENERAL INFORMATION:
; APPLICANT: Albertson, Marc C.
; APPLICANT: Fox, Tim W.
; APPLICANT: Carl, Garnaat W.
; APPLICANT: Huffman, Gary A.
; APPLICANT: Kendall, Timmy U.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION
; TITLE OF INVENTION: AND METHOD OF USING SAME
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.

```
STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.
STREET: Box 1000
CITY: Johnston
STATE: Iowa
COUNTRY: USA
ZIP: 50131
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/880,499
FILING DATE: CONCURRENTLY HEREWITH
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Sweeney, Patricia A.
REGISTRATION NUMBER: 32,733
REFERENCE/DOCKET NUMBER: 0578
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 248-4800
TELEFAX: (515) 248-4844
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1394 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-880-499-1
```

```
Query Match          4.2% Score 55.6; DB 3; Length 1394;
Best Local Similarity 55.8%; Pred. No. 0.00064;
Matches 106; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 761 TAGATTCTCGTCAAAAATCTTTCTGATTTTAAAGCTAGTTGGCAACCTGTTTC 820
DB 950 TACACAGTATGTATTTATTTAGGAAACATTCAGGCTAGCTGGCAACTATTTT 891
QY 821 TTTCAGAAATTTGATTTTTCAGAAAATAGTTATTTCTCTTATAAATAGAA 880
DB 890 TCTAAGTGTCTTATTTTATAAGAGAAATTAACCTAATTTTTTGAATAATCAAA 831
QY 881 AACCTTAGAAAATAGAGTTGCCAGACTAGCCCTAGATGTTTCCCATTAATTACA 940
DB 830 TTCTTTGAAGAAACAGGGTTCCTCAACTGCTTTAAAAAATCAGAAAGATTTTGA 771
QY 941 TCACTGTGTA 950
DB 770 CGAGAAATCTA 761
```

```
RESULT 10
US-08-880-499-2/c
Sequence 2, Application US/08880499
Patent No. 6037523
GENERAL INFORMATION:
APPLICANT: Albertson, Marc C.
APPLICANT: Fox, Tim W.
APPLICANT: Carl, Garraat W.
APPLICANT: Huffman, Gary A.
APPLICANT: Kendall, Jimmy L.
TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION
TITLE OF INVENTION: AND METHOD OF USING SAME
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESSES:
ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.
STREET: Box 1000
CITY: Johnston
STATE: Iowa
COUNTRY: USA
ZIP: 50131
```

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/880,499
FILING DATE: CONCURRENTLY HEREWITH
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Sweeney, Patricia A.
REGISTRATION NUMBER: 32,733
REFERENCE/DOCKET NUMBER: 0578
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 248-4800
TELEFAX: (515) 248-4844
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1394 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-880-499-2
```

```
Query Match          4.2% Score 55.6; DB 3; Length 1394;
Best Local Similarity 55.8%; Pred. No. 0.00064;
Matches 106; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 761 TAGATTCTCGTCAAAAATCTTTCTGATTTTAAAGCTAGTTGGCAACCTGTTTC 820
DB 950 TACACAGTATGTATTTATTTAGGAAACATTCAGGCTAGCTGGCAACTATTTT 891
QY 821 TTTCAGAAATTTGATTTTTCAGAAAATAGTTATTTCTCTTATAAATAGAA 880
DB 890 TCTAAGTGTCTTATTTTATAAGAGAAATTAACCTAATTTTTTGAATAATCAAA 831
QY 881 AACCTTAGAAAATAGAGTTGCCAGACTAGCCCTAGATGTTTCCCATTAATTACA 940
DB 830 TTCTTTGAAGAAACAGGGTTCCTCAACTGCTTTAAAAAATCAGAAAGATTTTGA 771
QY 941 TCACTGTGTA 950
DB 770 CGAGAAATCTA 761
```

```
RESULT 11
US-09-806-708B-23
Sequence 23, Application US/09806708B
Patent No. 6784342
GENERAL INFORMATION:
APPLICANT: The University of British Columbia
TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
FILE REFERENCE: 4810-58741
CURRENT APPLICATION NUMBER: US/09/806,708B
CURRENT FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: US 60/147,133
PRIOR FILING DATE: 1998-08-04
NUMBER OF SEQ ID NOS: 23
SOFTWARE: Patentin version 3.0
SEQ ID NO 23
LENGTH: 1055
TYPE: DNA
ORGANISM: Artificial sequence
FEATURE:
NAME/KEY: promoter
LOCATION: (1)..(1055)
OTHER INFORMATION: consensus sequence of A.c. and L.a. FAEI promoters
US-09-806-708B-23
```

```
Query Match          4.2% Score 55.4; DB 4; Length 1055;
Best Local Similarity 22.5%; Pred. No. 0.00065;
Matches 182; Conservative 181; Mismatches 430; Indels 15; Gaps 3;
```


Oy		292	CCACCAATTAATTATTTAGTAAATTAATTTAATTTTTGGAAAAATAACAACCTTATAC	351
Db		37	YCAANTGRCVYARBGKMTTAYMTATKMGITGMAMRTWRMAMKKTRKMCGSTLMMNNW	96
Oy		352	TTTTGTGAGGGCCTCAG-CATGATATTTTCGTTAGGGGCCGAATGCGAGACC	410
Db		97	TTCCTABRKMRGTBKWTTKTNNNATGBRMWTGYMTMNNGCSPTWTABRYKTRRMCYTAM	156
Oy		411	ATGCTTAGTGCTCACTATTGGCACCTAACCAACAAGATTAAAATAAACCAAATRA	470
Db		157	WYGASMAENASTRRRTTYTWRMCMCKRSBARATRGRARYRARVTMARREGTWKAMA	216
Oy		471	CCTAACCCA CTGGAAGCTCATCATGTAGT---TTAAGAACA CACTATTTAAACACAGA	527
Db		217	YAAMTMNNNNNNNAKAKCRKATTWGRBASNCCTTAGGTTTKRATCCMAVTYGAEMATKC	276
Oy		528	TCCTCTTAAAAAACAAGCATATTTGCAAGAAGACAATYANGTTACAGTTTACAAACATC	587
Db		277	WKWTXSAAMGMTVMNNNNNTTTTAKALMYAABRWMSATTTWMAAAWTSKWTYYGR	336
Oy		588	TAAGGCACCAATTAATATCGAAGCTAGCTATGACGTTCAAGATTTTCTTTTCATTC	647
Db		337	KTAANNNGTTCWTEWMAWMYCMWKKTKGTTTNNNGSRTYGMWTKKXAATTTAKANNC	396
Oy		648	TTGTTATTTTGTATGTTTATADACATTTCTCTCTTACATADAGATGATTTTCCT	707
Db		397	TTAAMKMKTCIMNNTTAAKATTYMATCYKWSHNTGSTRYABARTTWTAWTRRAYANN	456
Oy		708	CCGATTTTATAAATGACTATA-----AAGCATTTTATATAAGACACGCAT	756
Db		457	NNTKTTWACCTWTTYKRCCTTANNTPAAMYTXSANCTSHTRKTNKNCWRBAGSKTSMGRA	516
Oy		757	GTCGAGATTCCTCGTTCAAAAATCTTTCGATTTTTTAAAGCTAGTTTGGCAACCTCG	816
Db		517	YARAATWGTYKNNTAWAYCMTWYYYRAGAAMTAMTMTSATCYCAPAATTTAGTCAGAGS	576
Oy		817	TTTCTTTCAAGAATTTGATTTTTCACAAAAAATTAATTATTTCTCTTTATAAAT	876
Db		577	TAKGNNNNNNNNCCAACTCARMKCTPASAACMAANAATTCYCYAANNAATYMANNAATGW	636
Oy		877	AGAAAAACCTAGAAAAATPAGAGTTGCCAGACTAGCCCTAGAAATTTTTCCCAATAAT	936
Db		637	NATKTAATMTNNNNNNAGTWNNNNNNAKMASAATWYAAAAATATKYARTAANTTAAG	696
Oy		937	ACAATCACTGTGTATAATTATTTGGCCAGCCCCCAATTAATTTAAACGAAACTGAAT	996
Db		697	AYARPAATYTRANNACCTTTTNTTGGMTNTPAARGWANNNNNNNNNNNGACWA	756
Oy		997	CGAGCGAACAACAAATCTGAGCTAATTTCTAGATTAAGTAAAAAGAGAGAGAGAGAAG	1056
Db		757	WRTTATATANCGRNNNNNNNNNNAVATTTNTATTTTWWTRKUNNNNNNNNAAAYYGAAW	816
Oy		1057	AAATCAGTTTAAAGTCATGTCCTCGAG	1084
Db		817	KNNTMCWTCRAMKAMATGAATTTNAG	844
 RESULT 12 US-09-004-056-1/c Sequence 1, Application US/09004056A Patent No. 6566586 GENERAL INFORMATION: APPLICANT: Calgene LLC TITLE OF INVENTION: Plant Expansin Promoter Sequences FILE REFERENCE: 125 CURRENT APPLICATION NUMBER: US/09/004_056A CURRENT FILING DATE: 1998-01-07 EARLIER APPLICATION NUMBER: 60034914 EARLIER FILING DATE: 1997-07-01 NUMBER OF SEQ ID NOS: 1 SOFTWARE: PatencIn Ver. 2.0 SEQ ID NO 1				

```

; LENGTH: 2614
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (930)
; OTHER INFORMATION: unknown nucleotide
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (947)
; OTHER INFORMATION: unknown nucleotide
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (956)
; OTHER INFORMATION: unknown nucleotide
US-09-004-056-1

Query Match          4.2%; Score 55.2; DB 4; Length 2614;
Best Local Similarity 54.2%; Pred. No. 0.001;
Matches 155; Conservative 0; Mismatches 128; Indels 3; Gaps 2

Oy 611 AGTAAGCTATGACGCTGCAGATTTTCTTTTCATCTTGTAATTTGTATTTGTTTAA 670
Db 447 AGTTAGTTGGTATATCAATTTTTCATATTAATTTTAGTTTTATTTCTAATTTTA 388
Oy 671 TATACATTTTCTCTTCAATAGAGCATTTTCTCCGATTTTATATAAATGACTATTA 730
Db 387 TGTGACAAATGAACTTATTTATATATTTTAAATTTTTCATATAAAT-TTTAAG 329
Oy 731 AGCATTTTATATAGAGCAGCATGTCGTAGATTCGTTCCAAAATCTTCTGATTT 790
Db 328 TATTTTTCATATATATATTTTACGAGAAAACAATATTTTGAATACGAATTTTGAGATT 269
Oy 791 TTTTACAGCTAGTTGGCAACCCGTGTTCTTTCAGAAATTTTGATTTTTCAAAAAA 850
Db 268 TTTAAATAGTTAGATTTTATATTTCTTTATTTATATGAAATAT-TTATTTAGAAAAA 211
Oy 851 ATTAGTTATTTTCTCTTATATAAATAGAAAAACACTTAGAAAAATA 896
Db 210 ATTGAGTTAATGAAATTTTAAATTTTAAATATATATAAAAAATA 165

RESULT 13
US-09-640-173-53
; Sequence 53, Application US/09640173
; Patent No. 6613515
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: OVARIAN TUMOR SEQUENCES AND
; TITLE OF INVENTION: METHODS OF USE THEREFOR
; FILE REFERENCE: 210121.484C2
; CURRENT APPLICATION NUMBER: US/09/640.173
; CURRENT FILING DATE: 2000-08-15
; NUMBER OF SEQ ID NOS: 136
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 53
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(396)
; OTHER INFORMATION: n = A,T,C or G
US-09-640-173-53

Query Match          4.2%; Score 55; DB 4; Length 396;
Best Local Similarity 43.8%; Pred. No. 0.0056;
Matches 139; Conservative 0; Mismatches 178; Indels 0; Gaps 0;

Oy 632 TTTTCTTTTCATCTGTATATTTGTTATTTTATATATACATTTTCTTCTTACA 691
Db 11 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 70

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 07:10:11 ; Search time 853.804 Seconds
(without alignments)
10230.248 Million cell updates/sec

Title: US-10-713-381-1_COPY_1_1311

Perfect score: 1311
Sequence: 1 cccatgctgctctatgaaaaa.....cgcacatagcttcctcgaaac 1311

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7389322 seqs, 333128559 residues

Total number of hits satisfying chosen parameters: 14778644

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/prodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/prodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/prodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/prodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/prodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/prodata/2/pubpna/US09C_PUBCOMB.seq:*
13: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
16: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
17: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
18: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
19: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
20: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
21: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
22: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
23: /cgn2_6/prodata/2/pubpna/US11_PUBCOMB.seq:*
24: /cgn2_6/prodata/2/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1311	100.0	1394	20	US-10-713-381-1
2	1311	100.0	1394	20	US-10-713-381-2
3	160.6	12.3	255	20	US-10-713-381-9
4	146	11.1	158	20	US-10-713-381-3
5	80.8	6.2	1261	20	US-10-425-115-134230
6	74	5.6	320	18	US-10-425-114-23340
7	74	5.6	624	18	US-10-425-114-16264

8	73	5.6	1326	20	US-10-425-115-141826	Sequence 141826, A
9	72.8	5.6	2445	18	US-10-425-114-32493	Sequence 32493, A
10	72.8	5.6	2729	20	US-10-425-115-83293	Sequence 83293, A
11	72.4	5.5	1203	20	US-10-425-115-51470	Sequence 51470, A
12	71.8	5.5	928	18	US-10-425-114-17816	Sequence 17816, A
13	71.8	5.5	1123	20	US-10-425-115-142853	Sequence 142853, A
14	71.8	5.5	2537	18	US-10-425-114-31957	Sequence 31957, A
15	71.8	5.5	2537	20	US-10-425-115-52216	Sequence 52216, A
16	71.8	5.5	3607	18	US-10-425-114-31061	Sequence 31061, A
17	71.8	5.5	3691	20	US-10-425-115-52219	Sequence 52219, A
18	70.2	5.4	960	14	US-10-198-846-6381	Sequence 6381, A
19	69.8	5.3	13400	21	US-10-343-477A-63	Sequence 63, Appl
20	69.6	5.3	724	20	US-10-425-115-87756	Sequence 87756, A
21	69.2	5.3	527	18	US-10-425-114-4041	Sequence 4041, A
22	68.8	5.2	610	20	US-10-425-115-47165	Sequence 47165, A
23	68.4	5.2	1215	18	US-10-425-114-24656	Sequence 24656, A
24	68.4	5.2	2863	20	US-10-425-115-75310	Sequence 75310, A
25	67.8	5.2	512	20	US-10-425-115-49781	Sequence 49781, A
26	67.8	5.2	928	18	US-10-425-114-17816	Sequence 17816, A
27	67.8	5.2	1123	20	US-10-425-115-142853	Sequence 142853, A
28	67.4	5.1	1260	18	US-10-425-114-30881	Sequence 30881, A
29	67.4	5.1	1411	20	US-10-425-115-149304	Sequence 149304, A
30	67.2	5.1	2232	18	US-10-425-114-16607	Sequence 16607, A
31	67.2	5.1	2249	18	US-10-425-114-20264	Sequence 20264, A
32	66.4	5.1	6027	18	US-10-634-262-1	Sequence 1, Appl
33	66.4	5.1	6027	19	US-10-109-048-1145	Sequence 1145, A
34	66.4	5.1	6051	20	US-10-425-115-41589	Sequence 41589, A
35	66.2	5.0	439	9	US-09-960-352-1009	Sequence 1009, A
36	66	5.0	1376	18	US-10-425-114-2073	Sequence 2073, A
37	65.6	5.0	721	20	US-10-425-115-17795	Sequence 17795, A
38	65.6	5.0	2274	18	US-10-425-114-777	Sequence 777, A
39	65.6	5.0	2274	18	US-10-425-114-830	Sequence 830, A
40	65.6	5.0	2632	20	US-10-425-115-162957	Sequence 162957, A
41	65.6	5.0	8056	20	US-10-473-126-386	Sequence 386, A
42	65	5.0	1364	20	US-10-425-115-150820	Sequence 150820, A
43	64.8	4.9	833	20	US-10-425-115-30824	Sequence 30824, A
44	64.8	4.9	1838	20	US-10-425-115-21012	Sequence 21012, A
45	64.2	4.9	581	20	US-10-425-115-48324	Sequence 48324, A

ALIGNMENTS

RESULT 1
US-10-713-381-1
; Sequence 1, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OR INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/860,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-1
Query Match 100.0%; Score 1311; DB 20; Length 1394;
Best Local Similarity 100.0%; Pred. No. 5e-275;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCATGTCCTCTATGAAAAGATGATGATGCTCTATATCCGTTTCTTAGGTC 60

Db 1 CCAATGCTCTATGAAAAAGATGAGTCAATGTCTATATCCGTTTCTTAGGCTC 60
Qy 61 CTTCTTCTGCTTATTAAGTCTGAATCGGGGTTACAAAACTTCCAGGGTGAT 120
Db 61 CTTCTTCTGCTTATTAAGTCTGAATCGGGGTTACAAAACTTCCAGGGTGAT 120
Qy 121 CTCATGTTCCATTCCTCCACCTCGGGTGCACTTCTTGGAATCGGTGGTCCCAT 180
Db 121 CTCATGTTCCATTCCTCCACCTCGGGTGCACTTCTTGGAATCGGTGGTCCCAT 180
Qy 181 CTGACCGAGGCCCATGAGACACCTTTCGGACACCCATCAAGGAGCCCTTCGATGCCCCA 240
Db 181 CTGACCGAGGCCCATGAGACACCTTTCGGACACCCATCAAGGAGCCCTTCGATGCCCCA 240
Qy 241 CGAGAGCTATCGGGTGTGTATCCAGGGATATATGTCGCCCAATGCTCATTA 300
Db 241 CGAGAGCTATCGGGTGTGTATCCAGGGATATATGTCGCCCAATGCTCATTA 300
Qy 301 TTATTAATCTTGAATATTAATTTTGGAAAAATACAACTTATCTTTGTTGA 360
Db 301 TTATTAATCTTGAATATTAATTTTGGAAAAATACAACTTATCTTTGTTGA 360
Qy 361 GGGGCTCAGCATGATTTTCGCTTAGGGCCAGAAAATGCGAGGACCGCATGCTAGT 420
Db 361 GGGGCTCAGCATGATTTTCGCTTAGGGCCAGAAAATGCGAGGACCGCATGCTAGT 420
Qy 421 TCACATTTGGCATCCCGAGAACAAAGTTAAAAAATACCAAGTAACTATCACT 480
Db 421 TCACATTTGGCATCCCGAGAACAAAGTTAAAAAATACCAAGTAACTATCACT 480
Qy 481 CGAAGCTATCATGATGTTTAAAGAAACATCTATTTAAACACGATCCTTTAAAAA 540
Db 481 CGAAGCTATCATGATGTTTAAAGAAACATCTATTTAAACACGATCCTTTAAAAA 540
Qy 541 CAAGCATATTTGAGAAAGACAAATTAATGTTACAGTTTACAAACATCTAAGACGACAA 600
Db 541 CAAGCATATTTGAGAAAGACAAATTAATGTTACAGTTTACAAACATCTAAGACGACAA 600
Qy 601 TTATATCGAAGGTAGCTATGACGTTCCAGATTTTCTTTCAATCTGTTATTTGTT 660
Db 601 TTATATCGAAGGTAGCTATGACGTTCCAGATTTTCTTTCAATCTGTTATTTGTT 660
Qy 661 ATGTTTTATATACATTTCTCTCTTACATAGAGTATTTCTTCGATTTTAA 720
Db 661 ATGTTTTATATACATTTCTCTCTTACATAGAGTATTTCTTCGATTTTAA 720
Qy 721 ATGACTATAAAGTCAATTTTATATAAGACGCAATGCTAGATTTCTGTTCAAAAATC 780
Db 721 ATGACTATAAAGTCAATTTTATATAAGACGCAATGCTAGATTTCTGTTCAAAAATC 780
Qy 781 TTTCTATTTTAAAGAGTATGTCGCAACCTGTTCTTCAAGAAATTTGATTT 840
Db 781 TTTCTATTTTAAAGAGTATGTCGCAACCTGTTCTTCAAGAAATTTGATTT 840
Qy 841 TTCAAAAAAATAGTTATTTCTCTTATAAATAAGAAACACTTAGAAAAATAGAGT 900
Db 841 TTCAAAAAAATAGTTATTTCTCTTATAAATAAGAAACACTTAGAAAAATAGAGT 900
Qy 901 TGGCAGCTAGCCCTAGAAATGTTTCCCAATAAATACATCACTGTATATATTTG 960
Db 901 TGGCAGCTAGCCCTAGAAATGTTTCCCAATAAATACATCACTGTATATATTTG 960
Qy 961 GCGAGCCCATTAATATTAAACGAAACTGAAATCGAGCCGAACCAAACTGAGCTAT 1020
Db 961 GCGAGCCCATTAATATTAAACGAAACTGAAATCGAGCCGAACCAAACTGAGCTAT 1020
Qy 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAATCAGTTTAAAGTCACTGTCC 1080
Db 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAATCAGTTTAAAGTCACTGTCC 1080
Qy 1081 TGAATGTGGGTTTGGCAACATAGCACCGTAACTATAGTCAATAGTGTCCCTAGTCA 1140
Db 1081 TGAATGTGGGTTTGGCAACATAGCACCGTAACTATAGTCAATAGTGTCCCTAGTCA 1140

Db 1081 TGAATGTGGGTTTGGCAACATAGCACCGTAACTATAGTCAATAGTGTCCCTAGTCA 1140
Qy 1141 GGTGGGAGCTCTGTGTCACTCATGAGGATCTACATCTGTTTCAACCGTTGTC 1200
Db 1141 GGTGGGAGCTCTGTGTCACTCATGAGGATCTACATCTGTTTCAACCGTTGTC 1200
Qy 1201 TTGTTCCATTCGTCGAAGCCCTGCTATTCTGAAACCAAGAGATACCTACTCCCAACAT 1260
Db 1201 TTGTTCCATTCGTCGAAGCCCTGCTATTCTGAAACCAAGAGATACCTACTCCCAACAT 1260
Qy 1261 CCATCTTACTCATGCACTTCCATGCAACAGCAATATGTTCTGTAAC 1311
Db 1261 CCATCTTACTCATGCACTTCCATGCAACAGCAATATGTTCTGTAAC 1311

RESULT 2
US-10-713-381-2
; Sequence 2, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAAT, CARL W.
; APPLICANT: HUFEMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 576R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-2

Query Match 100.0%; Score 1311; DB 20; Length 1394;
Best Local Similarity 100.0%; Pred. No. 5e-275;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCAATGCTCTATGAAAAAGATGAGTCAATGTCTATATCCGTTTCTTAGGCTC 60
Db 1 CCAATGCTCTATGAAAAAGATGAGTCAATGTCTATATCCGTTTCTTAGGCTC 60
Qy 61 CTTCTTCTGCTTATTAAGTCTGAATCGGGGTTACAAAACTTCCAGGGTGAT 120
Db 61 CTTCTTCTGCTTATTAAGTCTGAATCGGGGTTACAAAACTTCCAGGGTGAT 120
Qy 121 CTCATGTTCCATTCCTCCACCTCGGGTGCACTTCTTGGAATCGGTGGTCCCAT 180
Db 121 CTCATGTTCCATTCCTCCACCTCGGGTGCACTTCTTGGAATCGGTGGTCCCAT 180
Qy 181 CTGACCGAGGCCCATGAGACACCTTTCGGACACCCATCAAGGAGCCCTTCGATGCCCCA 240
Db 181 CTGACCGAGGCCCATGAGACACCTTTCGGACACCCATCAAGGAGCCCTTCGATGCCCCA 240
Qy 241 CGAGAGCTATCGGGTGTGTATCCAGGGATATATGTCGCCCAATGCTCATTA 300
Db 241 CGAGAGCTATCGGGTGTGTATCCAGGGATATATGTCGCCCAATGCTCATTA 300
Qy 301 TTATTAATCTTGAATATTAATTTTGGAAAAATACAACTTATCTTTGTTGA 360
Db 301 TTATTAATCTTGAATATTAATTTTGGAAAAATACAACTTATCTTTGTTGA 360
Qy 361 GGGGCTCAGCATGATTTTCGCTTAGGGCCAGAAAATGCGAGGACCGCATGCTAGT 420
Db 361 GGGGCTCAGCATGATTTTCGCTTAGGGCCAGAAAATGCGAGGACCGCATGCTAGT 420
Qy 421 TCACATTTGGCATCCCGAGAACAAAGTTAAAAAATACCAAGTAACTATCACT 480
Db 421 TCACATTTGGCATCCCGAGAACAAAGTTAAAAAATACCAAGTAACTATCACT 480

```

Db      421  TCCACTTTGGCACTACCCAGAACAGATTTAAABAAAATACCAAAGTAACTAATCCACT 480
Qy      481  CGAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAACCAAGTCTCTTAAAAAA 540
Db      481  CGAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAACCAAGTCTCTTAAAAAA 540
Qy      541  CAAGCATATTTGCAAGAGACAAATTATGTATGATCTTAAACATCTAAGACGCAAA 600
Db      541  CAAGCATATTTGCAAGAGACAAATTATGTATGATCTTAAACATCTAAGACGCAAA 600
Qy      601  TTATATCGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Db      601  TTATATCGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Qy      661  ATTGTTTATATATATATATATATATATATATATATATATATATATATATATATAT 720
Db      661  ATTGTTTATATATATATATATATATATATATATATATATATATATATATATATAT 720
Qy      721  ATGACTATTAAGTATTTTATATATATATATATATATATATATATATATATATATAT 780
Db      721  ATGACTATTAAGTATTTTATATATATATATATATATATATATATATATATATATAT 780
Qy      781  TTTCTGATTTTATATATATATATATATATATATATATATATATATATATATATAT 840
Db      781  TTTCTGATTTTATATATATATATATATATATATATATATATATATATATATATAT 840
Qy      841  TTCAAAAAAATAGTTATTTTCTCTTATATATATATATATATATATATATATATAT 900
Db      841  TTCAAAAAAATAGTTATTTTCTCTTATATATATATATATATATATATATATATAT 900
Qy      901  TGGCAGACTAGCCCTGAAATGTTTCCCAATTAATTCATCTGCTGTAATTTATTTG 960
Db      901  TGGCAGACTAGCCCTGAAATGTTTCCCAATTAATTCATCTGCTGTAATTTATTTG 960
Qy      961  GCCAGCCCCATTAATTTATTAACCGAACTGAAATGAGCAAAACCAATCTGAGCTAT 1020
Db      961  GCCAGCCCCATTAATTTATTAACCGAACTGAAATGAGCAAAACCAATCTGAGCTAT 1020
Qy      1021  TTCTCTAGATTAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Db      1021  TTCTCTAGATTAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Qy      1081  TGAAGATGCGGTTTGGCAACGATAGCAGCAGTATCTATGCTCATAGGCTCTAGCTCA 1140
Db      1081  TGAAGATGCGGTTTGGCAACGATAGCAGCAGTATCTATGCTCATAGGCTCTAGCTCA 1140
Qy      1141  GGTTCGGAGCTCTCGTGTATCTCATGATGATGATGATGATGATGATGATGATGATG 1200
Db      1141  GGTTCGGAGCTCTCGTGTATCTCATGATGATGATGATGATGATGATGATGATGATG 1200
Qy      1201  TTGTTCCATGTCGCAAGCCTTGCTTATTTGAAACCAAGAGATACCTACCCAAACAT 1260
Db      1201  TTGTTCCATGTCGCAAGCCTTGCTTATTTGAAACCAAGAGATACCTACCCAAACAT 1260
Qy      1261  CCATCTTACATGCAACTTCCATGCAAAACGCAATATGTTTCTGAAC 1311
Db      1261  CCATCTTACATGCAACTTCCATGCAAAACGCAATATGTTTCTGAAC 1311

```

RESULT 3
US-10-713-381-9

```

; Sequence 9, Application US/10713381
; Publication No. US2004022131A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFMAN, GARY L.
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R

```

```

; CURRENT APPLICATION NUMBER: US/10/713,381
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 255
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-9

Query Match      12.3%; Score 160.6; DB 20; Length 255;
Best Local Similarity 97.6%; Pred. No. 6.1e-25;
Matches 163; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

RESULT 4
US-10-713-381-3

```

; Sequence 3, Application US/10713381
; Publication No. US2004022131A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFMAN, GARY L.
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 158
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-3

```

Query Match 11.1%; Score 146; DB 20; Length 158;

Best Local Similarity 99.4%; Pred. No. 7.3e-22;

Matches 157; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```

Qy      1155  CGTGCATCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1213
Db      1      CGTGCATCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
Qy      1214  CAAGCTTGCTTATTTCTGAACCAAGAGATACCTATCTCCAAACATCATCTTATCTCAT 1273
Db      61      CAAGCTTGCTTATTTCTGAACCAAGAGATACCTATCTCCAAACATCATCTTATCTCAT 120
Qy      1274  GCAACTTCATGCAAAACGCAATATGTTTCTGAAC 1311
Db      121      GCAACTTCATGCAAAACGCAATATGTTTCTGAAC 158

```

RESULT 5

US-10-425-115-134230

```
/ Sequence 134230, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(53222)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 134230
/ LENGTH: 1261
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MR14577_53903C.1
US-10-425-115-134230
```

```
Query Match          6.2%; Score 80.8; DB 20; Length 1261;
Best Local Similarity 78.2%; Pred. No. 3.1e-07;
Matches 97; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
```

```
QY 792 TTTAAGAGCTAGTTGGCAACCTGTTCTTCAAGAAATTTGATTTTCAAAAAA 851
    |||||
DB 999 TCTTAAGGCTAGTTGAAACCTTATTTTAAAGATTTTCTTTTAAAGTAA 1058
    |||||
QY 852 TTAGTTATTTTCTCTTATTAATAATGAAAAACCTTGAATAATGAGTTCAGACTAG 911
    |||||
DB 1059 TTAGTTATTTTCTCTTGAATAATGAAATCTCTGAAATAATGAGTTCAGACTAG 1118
    |||||
QY 912 CCCT 915
    |||||
DB 1119 CCCT 1122
    |||||
```

RESULT 6

```
US-10-425-114-23340
/ Sequence 23340, Application US/10425114
/ Publication No. US20040034888A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jindong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E.
/ APPLICANT: Tabaska, Jack E.
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(5313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 23340
/ LENGTH: 320
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: LIB3595-039-G8_F11
US-10-425-114-23340
```

```
Query Match          5.6%; Score 74; DB 18; Length 320;
Best Local Similarity 73.1%; Pred. No. 4.8e-06;
Matches 95; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
```

```
QY 791 TTTAAGAGCTAGTTGGCAACCTGTTCTTCAAGAAATTTGATTTTCAAAAAA 850
    |||||
DB 62 TTTTGGGCTAGTTGAAATCTCATTTTTCAGAGATTTTATTTTCCAAAGAA 121
    |||||
QY 851 ATTAGTTATTTCTCTTATTAATAATGAAAAACCTTGAATAATGAGTTCAGACTA 910
    |||||
```

```
DB 122 ATTAGTTATTTTCTCTTGAATAATGAAATCCCTGGAAATGAGTTCTAAACGA 181
QY 911 GCCCTAAT 920
    |||||
DB 182 GCCCTAAT 191
    |||||
```

RESULT 7

```
US-10-425-114-16264
/ Sequence 16264, Application US/10425114
/ Publication No. US20040034888A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jindong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E.
/ APPLICANT: Tabaska, Jack E.
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(5313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 16264
/ LENGTH: 624
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: LIB3062-023-G10_F11
US-10-425-114-16264
```

```
Query Match          5.6%; Score 74; DB 18; Length 624;
Best Local Similarity 73.1%; Pred. No. 6.7e-06;
Matches 95; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
```

```
QY 791 TTTAAGAGCTAGTTGGCAACCTGTTCTTCAAGAAATTTGATTTTCAAAAAA 850
    |||||
DB 374 TTTTGGGCTAGTTGAAATCTCATTTTTCAGAGATTTTATTTTCCAAAGAA 433
    |||||
QY 851 ATTAGTTATTTTCTCTTATTAATAATGAAAAACCTTGAATAATGAGTTCAGACTA 910
    |||||
DB 434 ATTAGTTATTTTCTCTTGAATAATGAAATCCCTGGAAATGAGTTCAGACTA 493
    |||||
QY 911 GCCCTAAT 920
    |||||
DB 494 GCCCTAAT 503
    |||||
```

RESULT 8

```
US-10-425-115-141826
/ Sequence 141826, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(53222)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 141826
/ LENGTH: 1326
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MR14577_60829C.1
US-10-425-115-141826
```

```
Query Match          5.6%; Score 73; DB 20; Length 1326;
```

```
Best Local Similarity 69.1%; Pred. No. 1.6e-05;
Matches 114; Conservative 0; Mismatches 50; Indels 1; Gaps 1;

QY 756 TGTGTGATGATCTGTTCACAAAATCTTGTGATTTTAAAGACTGTTGGCAACCT 815
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1090 TGAAGTGGTGGTGTGTGTATTTCACTTAACTGTAAGGACTAATTTGGGACCCA 1149

QY 816 GTTCTTCAAGAAATTTGATTTTTCACAAAATAATGTTATTTCTCTTATATAA 875
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1150 ATTT-TTCACATGGAATTTTCATTTTCTTAAGGAAAATTAATGTTCCCTTGAGAAA 1208

QY 876 TAGAAAACACTTAGAAAAATAGAGTTGCCAGACTAGCCCTAGAAAT 920
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1209 TAGGAATCCCTTAGAAAAAAATAGTTCCAAACTAGCCCTTAAT 1253

RESULT 9
US-10-425-114-32493
; Sequence 32493, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaka, Jack E
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ ID NOS: 73128
; LENGTH: 2445
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: clone ID: UC-ZMFLB73402B09_FLI
US-10-425-114-32493

Query Match 5.6%; Score 72.8; DB 18; Length 2445;
Best Local Similarity 73.6%; Pred. No. 2.3e-05;
Matches 106; Conservative 0; Mismatches 37; Indels 1; Gaps 1;

QY 792 TTAAAGCTAGTTGGCAACCTGTTCTT-TCAGAAATTTGATTTTTCACAAAAA 850
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1623 TCTTAGGGCTAGTTGAGAACCTTTTTCACAAAAGATTTTCATTTTTCACAAA 1682

QY 851 ATTGTTATTTTCTCTTTATATAATAGAAAACCTTAGAAAATAGATTGCCAGCTA 910
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1683 ATTAGTCATTTTCTCTTGAAGAAATAGAAATCCCTTAATAATAATAGTTTCAAACTA 1742

QY 911 GCCCTAGATGTTTCCCAATAA 934
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1743 GCCCTAATGTTTTCATGAA 1766

RESULT 10
US-10-425-115-83293/C
; Sequence 83293, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5322)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
```

```
; SEQ ID NO 83293
; LENGTH: 2729
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(2729)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: clone ID: MRT4577_175978C.1
US-10-425-115-83293

Query Match 5.6%; Score 72.8; DB 20; Length 2729;
Best Local Similarity 60.7%; Pred. No. 2.5e-05;
Matches 136; Conservative 0; Mismatches 87; Indels 1; Gaps 1;

QY 792 TTAAAGCTAGTTGGCAACCTGTTCTTTCACAAAATTTGATTTTTCACAAAAA 851
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2523 TTAAAGGCTAGTTTAGAA-CCTCATTTTTCCTAGAGATTTCTATTTTCCAGAAAA 2465

QY 852 TTAGTTATTTTCTCTTTATATAATAGAAAACCTTAGAAAATAGATTGCCAGCTAG 911
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2464 ATATTTCATTTCCCTTGAAAAATAGAAATACCTTGAAAAAATAGAGTTTCCAAACTAG 2405

QY 912 CCTAGATGTTTCCCAATTAATTAATCACTGTGTATTAATTTGGCCAGCCCAT 971
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2404 CTTCAAAATTTCTCTCAAGTATTTCTTCTTAGAAGTAAATGCTACATGGGAGCT 2345

QY 972 AAATTTTAAACCGAACTGAATCGACGAAACCAATCTGA 1015
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2344 TGAGTATCATGACTAAATGTAAAGGTAATAATCAATGTTA 2301

RESULT 11
US-10-425-115-51470/C
; Sequence 51470, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5322)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 51470
; LENGTH: 1203
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(1203)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: clone ID: MRT4577_146936C.1
US-10-425-115-51470

Query Match 5.5%; Score 72.4; DB 20; Length 1203;
Best Local Similarity 76.1%; Pred. No. 2e-05;
Matches 102; Conservative 0; Mismatches 31; Indels 1; Gaps 1;

QY 787 ATTTTAAAGCTAGTTGGCAACCTGTTCTTTCACAAAATTTGATTTTTCACAA 846
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1189 ACTTATTAAGGCTAGTGGGGA-CATATTTTCCAAAAGATTTCTATTTCTTAA 1131

QY 847 AAAAATTAATTTTCTCTTTATATAATAGAAAACCTTAGAAAATAGAGTTGCCAG 906
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1130 GAAATTAATTTTCTCTTGAAGAAATTAATAATCCGTAGAAAATAGAGTTCCAA 1071

QY 907 ACTAGCCCTAGAAAT 920
```


ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: MFT4577_147620C.1
US-10-425-115-52216

Query Match 5.5%; Score 71.8; DB 20; Length 2537;
Best Local Similarity 73.4%; Pred. No. 3.9e-05;
Matches 105; Conservative 0; Mismatches 37; Indels 1; Gaps 1;

QY	792	TTTAAGAGCTAGTTTGCAACCCCTGTTCTTCAAGAAATTTGATTTTTCAAAAAAA	851
Db	1799	TCTTAGGGCTAGTTTGAAACCCCT-TTTTCCCAAAAGATTTTCATTTTCAAGAAAA	1857
QY	852	TTAGTTATTTTCTCTTATATAAATAGAAAACTTAGAAAAATAGAGTTGCCAGACTAG	911
Db	1858	TTAGTTCATTTTCTTGAAGAAATAGAAATCCCTTAAAAAATAGGTTTTCAAACTAG	1917
QY	912	CCCTAGAAATGTTTCCCAATAA	934
Db	1918	CCCTTAATGTTTTCATGAA	1940

Search completed: September 15, 2005, 20:45:32
Job time : 854.804 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using bw model

Run on: September 15, 2005, 16:07:32 ; Search time 4088.83 Seconds
(without alignments)
12204.542 Million cell updates/sec

Title: US-10-713-381-1_COPY_1_1311

Perfect score: 1311
Sequence: 1 ccatggtgctctcctcgaataa.....cgacatcgtctcctcgaac 1311

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	908	69.3	963	9	CC656933	CC656933 OGMDO20TM
2	679	51.8	915	9	CG224225	CG224225 OGIAG08TV
3	419.8	32.0	687	9	CC656939	CC656939 OGMDO20TV
4	96	7.3	715	9	CG252571	CG252571 CGAB057C
5	92	7.0	967	9	CL235046	CL235046 ZMMBB057
6	88	6.7	814	9	CG048704	CG048704 PU1019TB
7	87.8	6.7	754	9	CG414922	CG414922 ZMMBB025
8	86.8	6.6	950	8	CC439901	CC439901 PUHRI15TB
9	86.6	6.6	652	8	CC384247	CC384247 PUHOC67TB
10	86.6	6.6	797	8	CC400575	CC400575 PUHLU61TD
11	86.6	6.6	820	8	CC400574	CC400574 PUHLU61TB
12	85.4	6.5	471	9	CG103452	CG103452 PUHBI9TB
13	85.2	6.5	765	9	CG082135	CG082135 PUFOX12TD
14	85.2	6.5	781	9	CG630219	CG630219 OGUCC53TV
15	85.2	6.5	815	9	CG349565	CG349565 OGOFI83TB
16	85.2	6.5	834	9	CC630210	CC630210 OGUCC53TV
17	85	6.5	789	8	CC433618	CC433618 PUHRI17TD
18	84.6	6.5	1092	8	CNS020K7	AL175696 Tetradon
19	84	6.4	793	8	BZ816381	BZ816381 PUPBA66TD
20	84	6.4	1078	9	CL997678	CL997678 ZMMBH001
21	83.8	6.4	530	9	CG201774	CG201774 PUICH24TB
22	83.6	6.4	722	9	CG333914	CG333914 OGOAD14TB
23	83.6	6.4	722	9	CG333929	CG333929 OGOAD14TV
24	83.6	6.4	781	9	CG034985	CG034985 PUIGR68TB

25	83.6	6.4	861	8	BZ797976	BZ797976 PUGBP28TB
26	83.6	6.4	865	8	CC430754	CC430754 PUHBP05TB
27	81.8	6.2	861	9	CG102092	CG102092 PUFYW94TD
28	81.8	6.2	947	8	CC435780	CC435780 PUHNS02TD
29	81.8	6.2	981	8	BZ784278	BZ784278 PUFHY09TD
30	81.8	6.2	1017	9	CC620594	CC620594 OGUCC66TV
31	81.6	6.2	649	9	CC613918	CC613918 OGIAG88TV
32	81.6	6.2	733	8	BZ778636	BZ778636 I102EL0.9
33	81.6	6.2	925	9	CG071791	CG071791 PUHBR02TB
34	81.6	6.2	960	8	BZ676889	BZ676889 PUHGI17TD
35	81.6	6.2	994	8	CC003943	CC003943 PUDJN65TD
36	81.6	6.2	1016	9	CL996481	CL996481 ZMMBH000
37	81.6	6.2	1022	9	CL984151	CL984151 ZMMBH000
38	81.4	6.2	999	8	CC385762	CC385762 PUHFX14TD
39	81.2	6.2	756	9	CG221653	CG221653 OGMW04TH
40	81	6.2	1101	9	CNS001T2	AL078714 Drosophila
41	80.6	6.1	699	8	BZ996930	BZ996930 PUGIR50TB
42	80.6	6.1	832	9	CC676387	CC676387 OGIAS18TB
43	80.6	6.1	839	9	CG246849	CG246849 OGMAL22TH
44	80.6	6.1	843	9	CG254365	CG254365 OGMF041TH
45	80.6	6.1	852	8	CC385881	CC385881 PUHMB02TD

ALIGNMENTS

RESULT 1	CC656933	963 bp	DNA	linear	GSS 19-JUN-2003
LOCUS	OGMDO20TM	ZM 0.7-1.5 kb	Zea mays	genomic clone	ZMMBMA0554D15,
DEFINITION	genomic survey sequence.				
ACCESSION	CC656933				
VERSION	CC656933.1	GI:32060225			
KEYWORDS	GSS.				
SOURCE	Zea mays				
ORGANISM	Zea mays				
REFERENCE	Whitelaw, C.A., Quackenbush, J., Van Aken, S., Uterback, T., Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.				
AUTHORS	Whitelaw, C.A., Quackenbush, J., Van Aken, S., Uterback, T., Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.				
TITLE	Whitelaw, C.A., Quackenbush, J., Van Aken, S., Uterback, T., Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.				
JOURNAL	Unpublished (2002)				
COMMENT	Other GSSs: OGMDO20TV Contact: Cathy Whitelaw TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA Tel: 301-838-5843 Fax: 301-838-0208 Email: whitelaw@tigr.org Seq primer: TR Class: sheared ends.				
FEATURES	Location/Qualifiers				
source	1..963 /organism="Zea mays" /mol_type="genomic DNA" /strain="B73" /db_xref="taxon:4577" /clone="ZMMBMA0554D15" /clone_lib="ZM_0.7-1.5_KB" /note="Vector: pBCK-1 Site 1: HincII, 0.7-1.5 kb methylation filtered genomic DNA library"				
ORIGIN					
Query Match	69.3%	Score 908	DB 9	Length 963	
Best Local Similarity	100.0%	Pred. No. 7.8e-180	Indels 0	Gaps 0	
Matches 908	Conservative 0	Mismatches 0	Indels 0	Gaps 0	
QY	404	ACCAGCAGTGTCTAGTGTCCATCTATGGCACTACCAAGCAAGATTTAAAAATTAACC	463		
DB	1	ACCAGCAGTGTCTAGTGTCCATCTATGGCACTACCAAGCAAGATTTAAAAATTAACC	60		

```
QY 464 AAGTAATCATCTCGAAGAGCTATCATGTATGTTTAAAGAAATCTATTTAAACC 523
DB 61 AAGTAATCATCTCGAAGAGCTATCATGTATGTTTAAAGAAATCTATTTAAACC 120
QY 524 AGCATCTCTTAAACAAAGCATATTTGCAAGAGACAAATATGTTTACAGTTTACAA 583
DB 121 AGCATCTCTTAAACAAAGCATATTTGCAAGAGACAAATATGTTTACAGTTTACAA 180
QY 584 CATCTAGAGCGCAATATATATGCAAGAGTATGACCTTGCAGATTTTCTTTTC 643
DB 181 CATCTAGAGCGCAATATATATGCAAGAGTATGACCTTGCAGATTTTCTTTTC 240
QY 644 ATCTGTATTTTGTATGTTTATATACATTTTCTCTTCAATAGAGTATTT 703
DB 241 ATCTGTATTTTGTATGTTTATATACATTTTCTCTTCAATAGAGTATTT 300
QY 704 TCTTCGATTTTAAATGACTATATAAGTCAATTTTATATAGACGACGATGCTAG 763
DB 301 TCTTCGATTTTAAATGACTATATAAGTCAATTTTATATAGACGACGATGCTAG 360
QY 764 ATCTGTCTCAAAATCTTCTGATTTTATAGAGTATTTGCAACCTGTTCTTT 823
DB 361 ATCTGTCTCAAAATCTTCTGATTTTATAGAGTATTTGCAACCTGTTCTTT 420
QY 824 CAAGAATTTTGAATTTTCAAAAAAATAGTTTATTTCTCTTATATAAGAAAC 883
DB 421 CAAGAATTTTGAATTTTCAAAAAAATAGTTTATTTCTTATATAAGAAAC 480
QY 884 ACTTGAATAATAGAGTGGCAGACTAGCCCTAGAAATGTTTTCCAAATAATTCATCA 943
DB 481 ACTTGAATAATAGAGTGGCAGACTAGCCCTAGAAATGTTTTCCAAATAATTCATCA 540
QY 944 CTGTGATATATATTTTGGCAGCCCATATAATTTTAAACGAAATGCGGCA 1003
DB 541 CTGTGATATATATTTTGGCAGCCCATATAATTTTAAACGAAATGCGGCA 600
QY 1004 AACCAATCTGAGCTATTTCTCTAGATTAAGTAAAGAGAGAGAGAGAAATTCAG 1063
DB 601 AACCAATCTGAGCTATTTCTCTAGATTAAGTAAAGAGAGAGAGAGAAATTCAG 660
QY 1064 TTTTAAAGTATGTCCTTGAAGATGTCGGTGTGGCAACGATGCCATCATAGCT 1123
DB 661 TTTTAAAGTATGTCCTTGAAGATGTCGGTGTGGCAACGATGCCATCATAGCT 720
QY 1124 CATAGAGCTAGTGGCAGCTGCTGTGATCTGACATGCAATGCGATCTATATG 1183
DB 721 CATAGAGCTAGTGGCAGCTGCTGTGATCTGACATGCAATGCGATCTATATG 780
QY 1184 TTGTTCAACCGTCTGTTGTCATGTCGAAGCCTTGCTATTTCTGAACCAAGAGAT 1243
DB 781 TTGTTCAACCGTCTGTTGTCATGTCGAAGCCTTGCTATTTCTGAACCAAGAGAT 840
QY 1244 ACTTACTCCCAACAAATCATCTTACTCATGCAACTTCATGCAACGCAATATGTT 1303
DB 841 ACTTACTCCCAACAAATCATCTTACTCATGCAACTTCATGCAACGCAATATGTT 900
QY 1304 TCGTGAAC 1311
DB 901 TCGTGAAC 908

RESULT 2
CG224225 915 bp DNA linear GSS 22-ANG-2003
DEFINITION OGIAG08TV_ZM_0.7.1.5_KB Zea mays genomic clone ZMMBma071B15,
CG224225 genomic survey sequence.
ACCESSION CG224225
VERSION CG224225.1 GI:34124113
KEYWORDS GSS-
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```

```
REFERENCE
1 (bases 1 to 915)
Whitehead, C.A., Quackenbush, J., Van Aken, S., Uterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Cleck, R.W., Numborg, A., Robbins, D. and Lake, N.
Consortium for Maize Genomics
Unpublished (2002)
COMMENT
Other GSSs: OGIAG08TV
Contact: Cathy Whitehead
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitehead@cigr.org
Seq primer: TP
Classes: sheared ends.
FEATURES
source
location/Qualifiers
1..915
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBma071B15"
/clone_lib="ZM 0.7.1.5_KB"
/notes="Vector: pBCK-; Site 1: HindII; 0.7-1.5 kb
methylation filtered genomic DNA library"
ORIGIN
Query Match 51.8%; Score 679; DB 9; Length 915;
Best Local Similarity 99.9%; Pred. No. 7.7e-132;
Matches 690; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 621 TGACCTTGAGATTTTCTTTTTCATTTCTGTTATTTTGTATGTTTATATATATTT 680
DB 1 TGACCTTGAGATTTTCTTTTTC-TTCTGTTATTTTGTATGTTTATATATATTT 59
QY 681 CTTCCTTACATAGAGATTTTCTCCGATTTTATATAATGACTATATAAGTCAATTTT 740
DB 60 CTTCCTTACATAGAGATTTTCTCCGATTTTATATAATGACTATATAAGTCAATTTT 119
QY 741 ATATAGAGCAGCATGTCGTAGATTCGTTCAAAAACTTTCTGATTTTATTAAGAC 800
DB 120 ATATAGAGCAGCATGTCGTAGATTCGTTCAAAAACTTTCTGATTTTATTAAGAC 179
QY 801 TAGTTGGCAACCTGTTCTTTCGAAGAATTTTGAATTTTCAAAAAATTAATTTAT 860
DB 180 TAGTTGGCAACCTGTTCTTTCGAAGAATTTTGAATTTTCAAAAAATTAATTTAT 239
QY 861 TTTCTTTTATTAATTAAGAAACACTTAGAAATATAGAGTTGCCAGACTAGCCTAGAAT 920
DB 240 TTTCTTTTATTAATTAAGAAACACTTAGAAATATAGAGTTGCCAGACTAGCCTAGAAT 299
QY 921 GTTTTCCCAATTAATTAATCAATCATCTGTATATATTTATTTGGCAGCCCATATATTT 980
DB 300 GTTTTCCCAATTAATTAATCAATCATCTGTATATATTTATTTGGCAGCCCATATATTT 359
QY 981 AAACCGAACTGAATTCAGAGGAAACCAATCTGAGCTATTTCTTAATTAATTAAG 1040
DB 360 AAACCGAACTGAATTCAGAGGAAACCAATCTGAGCTATTTCTTAATTAATTAAG 419
QY 1041 GGAGAGAGAGAGAAAGAAATAGTTTAAAGTCAATGTCCTGAGATGTCGGTTTGGCA 1100
DB 420 GGAGAGAGAGAGAAAGAAATAGTTTAAAGTCAATGTCCTGAGATGTCGGTTTGGCA 479
QY 1101 CGATAGCCACCGTAATCATAGCTCATAGTGCCTTCAAGTTGGGAGCTCTGTC 1160
DB 480 CGATAGCCACCGTAATCATAGCTCATAGTGCCTTCAAGTTGGGAGCTCTGTC 539
QY 1161 ATCTGACATGGCATCTCATGCTTGTCAACCGTTGCTTTGTCATCGTCAAGGCT 1220
DB 540 ATCTGACATGGCATCTCATGCTTGTCAACCGTTGCTTTGTCATCGTCAAGGCT 599
```

QY	1221	GGCCATTCTGAACCAAGAGTACTACTCCGAACATCATTTTCTCATGCACTT	1280
DB	600	TGCCATTCTGAACCAAGAGTACTACTCCGAACATCATTTCTCATGCACTT	659
QY	1281	CCATGCAACACGACATATGTTTCTCTGAC	1311
DB	660	CCATGCAACACGACATATGTTTCTCTGAC	690
RESULT 3			
CC656939/c			
LOCUS	CC656939	687 bp	DNA
DEFINITION	OGMDQ2TV_ZM_0.7.1.5_KB_Zea_mays_genomic_clone_ZMBA0554D15,		linear
ACCESSION	CC656939		Genomic survey sequence.
VERSION	CC656939.1		GI:32060231
KEYWORDS	GSS.		
SOURCE	Zea mays		
ORGANISM	Zea mays		
REFERENCE			
AUTHORS	Eukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta, Spermatophyta, Magnoliophyta, Liliopsida, Poales, Poaceae, PACAD clade; Panicoidae; Andropogoneae; Zea.		
	1 (bases 1 to 687)		
	Whitelaw,C.A., Quackenbush,J., Van Aken,S., Unterhach,T., Renick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nundberg,A., Robbins,D. and Lakey,N.		
TITLE	Consortium for Maize Genomics		
JOURNAL	Unpublished (2002)		
COMMENT	Other GSSs: OGMDQ20TM		
	Contact: Cathy Whitelaw		
	TIGR		
	9712 Medical Center Drive, Rockville, MD 20850, USA		
	Tel: 301-838-5843		
	Fax: 301-838-0208		
	Email: whitelaw@tigr.org		
	Seq primer: TP		
	Class: sheared ends.		
FEATURES			
Source	Location/Qualifiers		
	1..687		
	/organism="Zea mays"		
	/mol_type="genomic DNA"		
	/strain="873"		
	/db_xref="taxon:4577"		
	/clone="ZMBA0554D15"		
	/clone_11be="Zm_0.7.1.5_KB"		
	/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb		
	methylation filtered genomic DNA library"		
ORIGIN			
Query Match	32.0%	Score 419.8;	DB 9; Length 687;
Beet local similarity	99.5%;	Fred. No. 1.6e-77;	
Matches 421; Conservative	0; Mismatches 2; Indels 0; Gaps 0;		
QY	889	GAATAATAGAGTTGGCAGACATAGCCCTGATGTTTCCAAATAATTACAATCAGTGT	948
DB	687	GAATAATAGAGTTGGCAGACATAGCCCTGATGTTTCCAAATAATTACAATCAGTGT	628
QY	949	TATAATTATTTGGCCAGCCCATTAATTATTAAACCGAACTGAAATCGAGCGAAACA	1008
DB	627	TATAATTATTTGGCCAGCCCATTAATTATTAAACCGAACTGAAATCGAGCGAAACA	568
QY	1009	AATCGAGCTATTTCTCTAGATTTAGTAAAGGAGAGAGAGAGAAAGAAATCAGTTT	1068
DB	567	AATCGAGCTATTTCTCTAGATTTAGTAAAGGAGAGAGAGAGAAAGAAATCAGTTT	508
QY	1069	AGTCAATTGCTCCTGAGATGTCGGTGTGGCAACGATAGCAACGTAATCATAGCTCATAG	1128
DB	507	AGTCAATTGCTCCTGAGATGTCGGTGTGGCAACGATAGCAACGTAATCATAGCTCATAG	448
QY	1129	GTGCTTAGTCAAGTTGGCAGAGCTTCTGTGTCATCTCAATGAGATCACTACATCTGTT	1188
DB	447	GTGCTTAGTCAAGTTGGCAGAGCTTCTGTGTCATCTCAATGAGATCACTACATCTGTT	388

Qy	1189	CACCGTTCGTCCTGTTGTCATCGTCCAAAGCCTTGCTATTCTGAACCAAGAGATACCTA	1248
Db	387	CACCGTTCGTCCTGTTGTCATCGTCCAAAGCCTTGCTATTCTGAACCAAGAGATACCTA	328
Qy	1249	CTCCCAACATTCATCTTACTCATGCAACTTCCATGCAACGAGCAATATGTTTCTG	1308
Db	327	CTCCCAACATTCATCTTACTCATGCAACTTCCATGCAACGAGCAATATGTTTCTG	268
Qy	1309	AAC 1311	
Db	267	AAC 265	
RESULT 4			
LOCUS	CG252571/c	715 bp	DNA
DEFINITION	OG4BB05TC ZM_0.7_1.5_KB Zea mays genomic clone ZMMBma0809B10,	linear	GSS 25-AUG-2003
ACCESSION	CG252571		
VERSION	CG252571.1		
KEYWORDS	GSS.		
SOURCE	Zea mays		
ORGANISM	Zea mays		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.		
AUTHORS	1 (bases 1 to 715) Whitelaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T., Resnick,A., Frazer,C.M., Budiman,M.A., Bedell,J.A., Kohlfing,T., Citek,R.W., Nuneberg,A., Robbins,D. and Lakey,N.		
TITLE	Consortium for Maize Genomics		
JOURNAL	Unpublished (2002)		
COMMENT	Contact: Cathy Whitelaw TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA Tel: 301-838-5843 Fax: 301-838-0208 Email: whitelaw@tigr.org Seq primer: TF Class: sheared ends.		
FEATURES			
source	Location/Qualifiers		
	1..715		
	/organism="Zea mays"		
	/mol_type="genomic DNA"		
	/strain="B73"		
	/db_xref="taxon:4577"		
	/clone="ZMMBma0809B10"		
	/clone_1bp="ZM_0.7_1.5_KB"		
	/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb		
	methylation filtered genomic DNA library"		
ORIGIN			
Query Match	7.3%;	Score 96;	DB 9;
Best Local Similarity	80.1%;	Pred. No. 1.2e-09;	Length 715;
Matches 125;	Conservative % 0;	Mismatches 30;	Indels 1;
		Gaps 1;	
Qy	135	TTCTCCCACTTCGCTTCGACATTTCTTGATGTGCGTGTTCCTCATGACCAAGGCGCCA	194
Db	356	TCAACACACCCGCTCTTGCTTCCTTTCTTGATGTGCGGCTTCGCTGATCGAGGCCCA	297
Qy	195	TCAGACACCTTTCGGGACACCCATCAAGGCGCTTCGGATGAGCCACGAGACGTATCGGG	254
Db	296	TAAAGCACCTGTGCGAAGAACCCATTAAAGGCTTTTCGATGTGCCACGAGACAT-CGGGG	238
Qy	255	TCGTGTGATTCAGGGGATATATGTCTCCCAATC	290
Db	237	TCGTGTGATACCCGGGGATATCTATCCCCCAAC	202
RESULT 5			
LOCUS	CU235046	967 bp	DNA
DEFINITION	ZMMBB0575001r ZMMBB (HindIII) Zea mays genomic clone	linear	GSS 15-JAN-2004

Accession	Version	Source	Organism	Reference Authors	Title	Journal	Comment
ZMMBB0575001.3							genomic survey sequence.
CL235046							
CL235046.1	GI:40891729						
GSS.							
Zea mays							
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.							
1 (bases 1 to 967)							
Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C., Zohovetz,A.V., Fuks,G., Yu,Y., Wing,R. and Messing,J.							
Sequencing of the maize genome at PGIR (2003c)							
Unpublished (2003)							
Contact: Bharti,A.K.							
Dr.Joachim Messing's lab							
The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers University							
190 Frelinghuysen Road, Piscataway, NJ 08854, USA							
Tel: 732 445 3801							
Fax: 732 445 5735							
Email: bharti@waksman.rutgers.edu							
Seq primer: SP6							
Class: BAC ends							
High quality sequence start: 64.							
Location/Qualifiers							
1..967							
/organism="Zea mays"							
/mol_type="genomic DNA"							
/culivar="B73"							
/db_xref="taxon:4577"							
/clone="ZMMBB0575001"							
/lab_host="E. coli DH10B"							
/clone_11b="ZMMBB (HindIII)"							
/note="Vector: pUC19; Site_1: HindIII; Site_2: HindIII"							
ORIGIN							
Query Match	7.0%	Score 92,	DB 9,	Length 967,			
Best Local Similarity	78.6%	Pred: No. 8.5e-09;					
Matches 110,	Conservative 0,	Mismatches 30,	Indels 0,	Gaps 0,			
Oy	150	TGCACATCTTGTGATGTCGGTGTGTCATCTGACGAGGCCCATCAGACCTTTGG	209				
Db	827	TGCGCTTCTTGATACCGCGGTCCCGCTGTATTAGGCGCATCAGACCTTTGG	886				
Oy	210	GACACCCATCAAGGGCTTTCGATGAGCCACGAGACGATTCGGTGTGATTCAGG	269				
Db	887	GTGGCCCATCAAGTGCCTTCGGATGATGCCACGAGGATATCCGGCCGTGTGACTGG	946				
Oy	270	GGATATATCTCCCAACT	289				
Db	947	GGATATCTGTCCCACTAGT	966				
RESULT 6							
CG048704							
LOCUS	CG048704	814 bp	DNA	linear	GSS 19-AUG-2003		
DEFINITION	PULL019TB ZM.0.6.1.0_KB Zea mays genomic clone ZMMBR0611C13,						
ACCESSION	CG048704						
VERSION	CG048704.1	GI:33920884					
KEYWORDS	GSS.						
SOURCE	Zea mays						
ORGANISM	Zea mays						
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.						
AUTHORS	Whitelaw,C.A., Quackenbush,J., Van Aken,S., Unterback,T., Resnick,A., Frazer,C.M., Yuan,Y., San Miguel,P., Ma,J. and Benneken,J.						
TITLE	Maize Genomics Consortium						
JOURNAL	Unpublished (2003)						

COMMENT	OTHER GSSs: PUI1019TD
	Contact: Cathy Whiteleaw
	7172 Medical Center Drive, Rockville, MD 20850, USA
	Tel: 301-838-5843
	Fax: 301-838-0208
	Email: whiteleaw@tigr.org
	Seq primer: TR
	Class: sheared ends.
FEATURES	Location/Qualifiers
SOURCE	1. 814
	/organism="Zea mays"
	/mol_type="genomic DNA"
	/strain="B73"
	/db_xref="taxon:4577"
	/clone_lib="ZMMB0611C13"
	/clone_lib="ZM_0.6-1.0_KB"
	/note="Vector: PCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
	Cot selected genomic DNA library"
ORIGIN	
Query Match	6.7%; Score 88; DB 9; Length 814;
Best Local Similarity	62.7%; Pred. No. 5.8e-08; Mismatches 90; Indels 1; Gaps 1;
Matches 153; Conservative 0;	
Qy	713 TTTATATAATGACTATTAAGTCATTTTATATATAAGACGACGATGCTGATCTCGTT 772
Db	487 TTTTAGAGAGCGTTGGTATGAGATGGAAGATATAGAGAAATATCTTTAGAGAAATGTTA 546
Qy	773 CAAAAATCTTCGATTTTATTAAGACTATGTTGGCAACCGCTGTTCTTCAAGAATT 832
Db	547 TAAAGACAGAAACATTTCTTAAGAGCTAGTTGGTAACCCATTTTTCAAAGATT 606
Qy	833 TTGATTTTTCACAAA-AAAATGATTTATTTCTCTTATTAATAGAAACACTTAGAA 891
Db	607 TTCATTCTCAGAGGAGAAATAGTTATTTCCCTTTGTAATAGTATATCATCGGAA 666
Qy	892 AAATGAGTTCCAGACTAGCCCTTAGATGTTTCCCAATTAATTAACATCACTGTAT 951
Db	667 AAATAGAGTTCTCAAACTAGCCCTTAGAGTTATATCCGATTAATCCCTATCTCATTTCT 726
Qy	952 AATT 955
Db	727 ATTT 730
RESULT 7	
CG414922/c	754 bp DNA 1linear GSS 08-SEP-2003
LOCUS	ZMMBB0250L09.r ZMMBB Zea mays genomic clone ZMMBB0250L09 3',
DEFINITION	genomic survey sequence.
ACCESSION	CG414922
VERSION	CG414922.1 GI:34505144
KEYWORDS	GSS.
SOURCE	Zea mays
ORGANISM	Zea mays
	Eukaryote; Viridiplantae; Streptophyte; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE	1 (bases 1 to 754)
AUTHORS	Yu, Y., Kim, H.R., Hatfield, J., Soderlund, C., Bharti, A.K., Messing, J. and Wing, R.
	Sequencing of the maize genome
	Unpublished (2003)
TITLE	Contact: Rod Wing
JOURNAL	Arizona Genomics Institute
COMMENT	University of Arizona
	Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
	85721-0088, USA
	Tel: 520 626 3967
	Fax: 520 621 9288
	Email: http://genome.arizona.edu
	PCR Primers

```

/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMB75.19D06"
/clone_1bp="2M 0.6 1.0 KB"
/note="Vector: PCR4-TOP0; Site_1:
Cot selected genomic DNA library"

```

834 TGATTTTTCAAAAAATTAGTTTATTTCTCTTATATAATGAAAAACATTAGAAA 893

ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 471)
 WhiteLaw, C.A., Quackenbush, J., Van Aken, S., Utecherback, T., Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and Bennetzen, J.
 Maize Genomics Consortium
 Unpublished (2003)
 Other_GSSs: PUB019TD
 Contact: Cathy WhiteLaw
 TIGR

TITLE
 JOURNAL
 COMMENT

FEATURES
 Source
 Location/Qualifiers
 1..471
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone_id="ZM0628D13"
 /note="Vector: pCR4-TOP0; Site 1: EcoRI; 0.6-1.0 kb high Cor selected genomic DNA library"

ORIGIN

Query Match 6.5%; Score 85.4; DB 9; Length 471;
 Best Local Similarity 69.5%; Pred. No. 2e-07;
 Matches 116; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 789 TTTTAAAGACGATGTTGGCAACCGTCTTCAAGAAATTTGATTTTCAAAA 848
 216 TTTTCTTAGAGCTAGTTGGAAATCCATTTTCCAGGATTTACATTTTTCAGGG 275
 QY 849 AAATAGTTATTTCTCTTATTAATAAGAAAACTTGAATAATAGATTGCCAC 908
 276 AAATAGTTATTTCTCTTATTAATAAGAAAACTTGAATAATAGATTGCCAC 335
 Db
 QY 909 TAGCCCTAGATGTTTCCCAATTAATTAATACATCTGTATTAATT 955
 336 TAGCGCTTAGACTATTTCTTAATAAGTAAGTCAATGAATTAAGATT 382
 Db

RESULT 13
 CG082135/c 765 bp DNA linear GSS 20-AUG-2003
 LOCUS PUF0X12TD ZM 0.6_1.0 KB Zea mays genomic clone ZMMBTA0712B24,
 DEFINITION genomic survey sequence.
 ACCESSION CG082135
 VERSION CG082135.1 GI:33964429
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 765)
 WhiteLaw, C.A., Quackenbush, J., Van Aken, S., Utecherback, T., Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and Bennetzen, J.
 Maize Genomics Consortium
 Unpublished (2003)
 Other_GSSs: PUF0X12TD
 Contact: Cathy WhiteLaw
 TIGR

TITLE
 JOURNAL
 COMMENT

FEATURES
 Source
 Location/Qualifiers
 1..761
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone_id="ZM0628D13"
 /note="Vector: pCR4-TOP0; Site 1: EcoRI; 0.6-1.0 kb high Cor selected genomic DNA library"

ORIGIN

Query Match 6.5%; Score 85.2; DB 9; Length 765;
 Best Local Similarity 71.8%; Pred. No. 2e-07;
 Matches 125; Conservative 0; Mismatches 48; Indels 1; Gaps 1;

QY 787 ATTTTAAAGACGATGTTGGCAACCGTCTTCAAGAAATTTGATTTTCAAA 846
 255 ATCTTCTATGACGATGTTGGCAACCTATTT-CTCAGCGGATTTTATTTTCAAT 197
 Db
 QY 847 AAAAATAGTTATTTCTCTTATTAATAAGAAAACTTGAATAATAGATTGCCAC 906
 196 GAAATATGTTATTTCTCTTGAATAATTAATCACTTGGAATAATAGATTCCAA 137
 Db
 QY 907 ACTAGCCCTAGAAATGTTTCCCAATTAATTAATCAATCACTGTATTAATTG 960
 136 ACTAGCCCTAATGATTTTATTAATAAGAAAAATGTCATTTTCATTG 83
 Db

RESULT 14
 CC630219 781 bp DNA linear GSS 19-JUN-2003
 LOCUS OGUG53TV ZM 0.7_1.5 KB Zea mays genomic clone ZMMBTA0406J09,
 DEFINITION genomic survey sequence.
 ACCESSION CC630219
 VERSION CC630219.1 GI:32003354
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 781)
 WhiteLaw, C.A., Quackenbush, J., Van Aken, S., Utecherback, T., Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
 Consortium for Maize Genomics
 Unpublished (2002)
 Other_GSSs: OGUG53TV
 Contact: Cathy WhiteLaw
 TIGR

TITLE
 JOURNAL
 COMMENT

FEATURES
 Source
 Location/Qualifiers
 1..781
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone_id="ZM0628D13"
 /note="Vector: pCR4-TOP0; Site 1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"

ORIGIN

Query Match 6.5%; Score 85.2; DB 9; Length 781;

Best Local Similarity 71.8%; Pred. No. 2.2e-07;
Matches 125; Conservative 0; Mismatches 48; Indels 1; Gaps 1;

```
QY 787 ATTTTAAAGTACTGTTGGCAACCTGTTCTTCAAGAAATTTGATTTTTCAAA 846
    |||||
DB 665 ATCTTCTAAATGACTAGTTGGAACTTATTT-TCACAGGATTTTATTTTTCAAAT 607
    |||||
QY 847 AAAAATTAAGTTTCTCTTATATAAATAGAAAACATTAGAAAATAGAGTTGCCAG 906
    |||||
DB 606 GAAATTAAGTTTCTCTTATAGAAAATATAATCATTGAGAAAATATAGTTCCAA 547
    |||||
QY 907 ACTAGCCCTAAGAAATGTTTCCCAATAAATTACATCATGTGTATATTTG 960
    |||||
DB 546 ACTAGCCCTAAGTATTTTATAAAATAGAGAAAATGCTATATTTTCATTG 493
    |||||
```

RESULT 15
CG349565/c 815 bp DNA linear GSS 26-AUG-2003

LOCUS OGOF183TH ZM 0.7.1.5 KB Zea mays genomic clone ZMMBMA0705N21,
DEFINITION genomic survey sequence.

ACCESSION CG349565
VERSION CG349565.1 GI:34266831
KEYWORDS GSS.

SOURCE Zea mays
ORGANISM Zea mays

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 815)

AUTHORS WhiteIaw, C.A., Quackenbush, J., Van Aken, S., Utecherback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunez, A., Robbins, D. and Lakey, N.

JOURNAL Consortium for Maize Genomics
COMMENT unpublished (2002)
Other_GSSB: OGOF183TV
Contact: Cathy WhiteIaw

TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208

Email: whiteIaw@tigr.org
Seq primer: TR
Class: sheared ends.

FEATURES
source location/Qualifiers
1..815

/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0705N21"
/note="Vector: pBCK-; Site_1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match 6.5%; Score 85.2; DB 9; Length 815;
Best Local Similarity 71.8%; Pred. No. 2.2e-07;

Matches 125; Conservative 0; Mismatches 48; Indels 1; Gaps 1;

```
QY 787 ATTTTAAAGTACTGTTGGCAACCTGTTCTTCAAGAAATTTGATTTTTCAAA 846
    |||||
DB 173 ATCTTCTAATGACTAGTTGGAACTTATTT-TCACAGGATTTTATTTTTCAAAT 115
    |||||
QY 847 AAAAATTAAGTTTCTCTTATATAAATAGAAAACATTAGAAAATAGAGTTGCCAG 906
    |||||
DB 114 GAAATTAAGTTTCTCTTATAGAAAATATAATCATTGAGAAAATATAGTTCCAA 55
    |||||
QY 907 ACTAGCCCTAAGTATTTTCCCAATAAATTACATCATGTGTATATTTG 960
    |||||
DB 54 ACTAGCCCTAAGTATTTTATAAAATAGAGAAAATGCTATATTTTCATTG 1
    |||||
```

Search completed: September 16, 2005, 08:08:18
Job time : 4089.83 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using bw model

Run on: September 16, 2005, 08:14:25 ; Search time 4104 Seconds

(without alignments)
16458.716 Million cell updates/sec

Title: US-10-713-381-2

Perfect score: 1394

Sequence: 1 ccattggtgctctcctatgaaga.....cttttcgctccaccaccatg 1394

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_mt:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1394	100.0	1394	6	AX224395	AX224395 Sequence
2	1394	100.0	1394	6	BD062177	BD062177 Male t18s
3	1311	94.0	1394	6	AX224394	AX224394 Sequence
4	1311	94.0	1394	6	BD062176	BD062176 Male t18s
5	1311	94.0	3343	8	AF360356	AF360356 Zea mays
6	240	17.2	255	6	AX224402	AX224402 Sequence
7	111	8.0	158	6	AX224396	AX224396 Sequence
8	50	3.6	50	6	AX224398	AX224398 Sequence
9	40	2.9	40	6	AX224399	AX224399 Sequence
10	30	2.2	30	6	AX224397	AX224397 Sequence
11	25	1.8	114817	2	AC149836	AC149836 Zea mays
12	24	1.7	10384	1	AE010607	AE010607 Purobacte
13	24	1.7	184509	2	AC113237	AC113237 Canis fam
14	24	1.7	212621	2	AC098210	AC098210 Rattus no
15	24	1.7	213753	2	AC110715	AC110715 Rattus no
16	24	1.7	239424	2	AC118504	AC118504 Rattus no
17	24	1.7	245072	2	CR381551	CR381551 Danio rer
18	24	1.7	252679	2	AC118394	AC118394 Rattus no
19	23	1.6	849	6	AX384588	AX384588 Sequence

20	23	1.6	93615	9	AC004950	AC004950 Homo sapi
21	23	1.6	106759	9	AL357079	AL357079 Human DNA
22	23	1.6	114214	8	AC125479	AC125479 Medicago
23	23	1.6	120180	5	BX323592	BX323592 Zebrafish
24	23	1.6	132698	2	AC005051	AC005051 Homo sapi
25	23	1.6	169876	10	AC124433	AC124433 Mus muscu
26	23	1.6	187002	2	AC134483	AC134483 Rattus no
27	23	1.6	204724	10	AC131912	AC131912 Mus muscu
28	23	1.6	206924	10	AC087780	AC087780 Mus muscu
29	23	1.6	228736	2	AC121614	AC121614 Rattus no
30	23	1.6	245121	2	AC126482	AC126482 Rattus no
31	23	1.6	255925	2	AC094360	AC094360 Rattus no
32	23	1.6	259743	2	AC113965	AC113965 Mus muscu
33	22	1.6	939	3	CG157915	CG157915 Crabeoetr
34	22	1.6	1154	3	AY617812	AY617812 Steirkeil
35	22	1.6	28168	3	CEY3789A	AL032625 Caenorhab
36	22	1.6	93132	10	AC084402	AC084402 Mus muscu
37	22	1.6	128625	8	AC147499	AC147499 Medicago
38	22	1.6	128907	9	AC105445	AC105445 Homo sapi
39	22	1.6	135726	8	AC147498	AC147498 Medicago
40	22	1.6	137477	5	BX004882	BX004882 Zebrafish
41	22	1.6	158833	8	AC134926	AC134926 Oryza sat
42	22	1.6	170059	9	AC026780	AC026780 Homo sapi
43	22	1.6	172117	9	AC024890	AC024890 Homo sapi
44	22	1.6	178252	2	AL356386	AL356386 Homo sapi
45	22	1.6	179671	5	AL929022	AL929022 Zebrafish

ALIGNMENTS

RESULT 1
AX224395
LOCUS AX224395
DEFINITION Sequence 2 From Patent W00160997.
ACCESSION AX224395
VERSION AX224395.1 GI:15554637
KEYWORDS
SOURCE
ORGANISM Zea mays
Zea mays
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
AUTHORS Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
TITLE Male tissue-preferred regulatory region and method of using same
JOURNAL Patent: WO 0160997-A 2 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)

FEATURES
source
1..1394
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN

Query Match 100.0%; Score 1394; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCATGGTGTCTCATGAAAAAGATGAGTCAATGTGTCATATCCGTTTCTTAAAGGTC 60
1 CCAATGATGCTCTTATGAAAAGATGAGTCAATGTGTCATATCCGTTTCTTAAAGGTC 60
DB 1 CCAATGATGCTCTTATGAAAAGATGAGTCAATGTGTCATATCCGTTTCTTAAAGGTC 60
QY 61 CTTCTTCTGCTTATATGACTGATCGGGGTTACAAAACTTCACGGGTGATGAT 120
61 CTTCTTCTGCTTATATGACTGATCGGGGTTACAAAACTTCACGGGTGATGAT 120
DB 61 CTTCTTCTGCTTATATGACTGATCGGGGTTACAAAACTTCACGGGTGATGAT 120
QY 121 CTCGATGTTCCACTTTCACCTCGCGTTGCAATTTCTTGAGATGCGGTGTTCCAT 180
121 CTCGATGTTCCACTTTCACCTCGCGTTGCAATTTCTTGAGATGCGGTGTTCCAT 180
DB 121 CTCGATGTTCCACTTTCACCTCGCGTTGCAATTTCTTGAGATGCGGTGTTCCAT 180
QY 181 CTGACCGAGCCCATGACACCTTGCGGAGACCCATGACAGGCGCTTGAGTGGCCCA 240
181 CTGACCGAGCCCATGACACCTTGCGGAGACCCATGACAGGCGCTTGAGTGGCCCA 240

Db 181 CTGACCGAGGCCCATCAGACACCTTTGCGGACACCATCAAGGCGCTTTGAGTGCCCA 240
 QY 241 CGAGAGTATCGGGTGTGTGATCCAGGGGATATATGTCCTCCCAAAATGCTACCTTAA 300
 Db 241 CGAGAGTATCGGGTGTGTGATCCAGGGGATATATGTCCTCCCAAAATGCTACCTTAA 300
 QY 301 TTATTTATCTTTAGATATTTATTTTGAAGAAAATACAAATCTTATCTTTGTGA 360
 Db 301 TTATTTATCTTTAGATATTTATTTTGAAGAAAATACAAATCTTATCTTTGTGA 360
 QY 361 GGGCTCAGACATAGATTTTCGCTTAGGGCCAGAAATCGAGAGACCAAGCATGTCTAGTG 420
 Db 361 GGGCTCAGACATAGATTTTCGCTTAGGGCCAGAAATCGAGAGACCAAGCATGTCTAGTG 420
 QY 421 TCACATTTGGGCTACCCGAGACAGATTTAAATAATACCAAGTAACTAACTCACT 480
 Db 421 TCACATTTGGGCTACCCGAGACAGATTTAAATAATACCAAGTAACTAACTCACT 480
 QY 481 CGAAGCTATCATGTATGTTTTAAAGAAACATCTAATTAACACGATCCTTTAAATAA 540
 Db 481 CGAAGCTATCATGTATGTTTTAAAGAAACATCTAATTAACACGATCCTTTAAATAA 540
 QY 541 CAAGCATATTTGCAAAAGACAAATTTAGTTCAGATTTACAAACATCTAAAGCGACAA 600
 Db 541 CAAGCATATTTGCAAAAGACAAATTTAGTTCAGATTTACAAACATCTAAAGCGACAA 600
 QY 601 TTATATCGAAAGGTAGCTATGACGTTCAGATTTTCTTTGATCTTGTATTTTGT 660
 Db 601 TTATATCGAAAGGTAGCTATGACGTTCAGATTTTCTTTGATCTTGTATTTTGT 660
 QY 661 ATTGTTTTATATACATTTTCTTCTTACATAGATGATTTTCTTCCGATTTTAA 720
 Db 661 ATTGTTTTATATACATTTTCTTCTTACATAGATGATTTTCTTCCGATTTTAA 720
 QY 721 ATGACTATTAAGCTATTTTATATTAAGAGACGCAATGCTAGATTTCTGTTCAAAAATC 780
 Db 721 ATGACTATTAAGCTATTTTATATTAAGAGACGCAATGCTAGATTTCTGTTCAAAAATC 780
 QY 781 TTTCTGATTTTAAAGCTAGTTTGGCAACCTGTCTTCAAGAAATTTGATTT 840
 Db 781 TTTCTGATTTTAAAGCTAGTTTGGCAACCTGTCTTCAAGAAATTTGATTT 840
 QY 841 TTCAAAAAAATAGTTTATTTCTCTTATTAATAAGAAACATCTAGAAAAATGAGT 900
 Db 841 TTCAAAAAAATAGTTTATTTCTCTTATTAATAAGAAACATCTAGAAAAATGAGT 900
 QY 901 TGGCAACTAGCCCTGGAATGTTTTCCCAATAATTCATCTGCTGATTAATTTTG 960
 Db 901 TGGCAACTAGCCCTGGAATGTTTTCCCAATAATTCATCTGCTGATTAATTTTG 960
 QY 961 GCCAGCCCCATTAATTTTAAACGAAACTGAAATCGAGCGAAACCAATCTGACTAT 1020
 Db 961 GCCAGCCCCATTAATTTTAAACGAAACTGAAATCGAGCGAAACCAATCTGACTAT 1020
 QY 1021 TTCTCAGATTAGTAAAGGAGAGAGAGAGAGAAATCACTTTTAAAGCTATTTGCC 1080
 Db 1021 TTCTCAGATTAGTAAAGGAGAGAGAGAGAGAAATCACTTTTAAAGCTATTTGCC 1080
 QY 1081 TGAAGATGTCGGGTTTGGCAACATAGCCGTAATCATAGCTCATAGGTGCTTACTGA 1140
 Db 1081 TGAAGATGTCGGGTTTGGCAACATAGCCGTAATCATAGGTGCTTACTGA 1140
 QY 1141 GGTTCGGAGCTCTGCTGTCATCTCATGTGGCATATCTACATGCTTTTCAACCGTTGTC 1200
 Db 1141 GGTTCGGAGCTCTGCTGTCATCTCATGTGGCATATCTACATGCTTTTCAACCGTTGTC 1200
 QY 1201 TTGTTTCATCTGTCAGAGCTTGGCTATTTCTGAACCAAGAGATACCTACTCCCAACAT 1260
 Db 1201 TTGTTTCATCTGTCAGAGCTTGGCTATTTCTGAACCAAGAGATACCTACTCCCAACAT 1260
 QY 1261 CCATCTTACTCATGCAACTTCCATGCAACACGACATATGTTTCTGAAACGATCTATT 1320
 Db 1261 CCATCTTACTCATGCAACTTCCATGCAACACGACATATGTTTCTGAAACGATCTATT 1320

QY 1321 AAAGATCACAACAGGTAGGCTTCTCCGCTAGCTTCCCTCTGCTGCGGATCTTTT 1380
 Db 1321 AAAGATCACAACAGGTAGGCTTCTCCGCTAGCTTCCCTCTGCTGCGGATCTTTT 1380
 QY 1381 CGTCCACACCATG 1394
 Db 1381 CGTCCACACCATG 1394
 RESULT 2
 BD062177 1394 bp DNA linear PAT 27-AUG-2002
 LOCUS BD062177
 DEFINITION Male tissue-preferred regulatory region and method of using same.
 ACCESSION BD062177.1 GI:22607782
 VERSION BD062177.1
 KEYWORDS JP 2001520523-A/2.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1394)
 Albertsen, M.C., Fox, T.W., Garmaat, C.W., Huffman, G.A. and Kendall, T.L.
 TITLE Male tissue-preferred regulatory region and method of using same
 JOURNAL Patent: JP 2001520523-A 2 30-OCT-2001;
 PIONEER HI BRED INTERNATIONAL INC
 PN JP 2001520523-A/2
 PD 30-OCT-2001
 PF 19-JUN-1998 JP 1999504910
 PR 23-JUN-1997 US 08/880499
 PI MARC C ALBERTSEN, TIMOTHY W FOX, CARL W GARMAT, GARY A HUFFMAN,
 PI TIMMY L KENDALL
 PC C12N15/82, C12N15/29, C12N9/24, C12N9/22, C12N9/10, C12N9/00 PC
 PC C07K14/34, C12Q1/68,
 PC A01H5/00
 CC Strandedness: Single;
 CC Topology: Linear;
 FH Key location/Qualifiers.
 FEATURES
 source location/Qualifiers
 1..1394
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Query Match 100.0%; Score 1394; DB 6; Length 1394;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CGATGTCCTCTAGAAAAAGATAGTACATGTCATTCGCTTTCTTAGGGTCC 60
 Db 1 CGATGTCCTCTAGAAAAAGATAGTACATGTCATTCGCTTTCTTAGGGTCC 60
 QY 61 CTTCTTGCCTTATTAAGTGAATCGGGTTTCAAAAACTTCAAGGTCATGAT 120
 Db 61 CTTCTTGCCTTATTAAGTGAATCGGGTTTCAAAAACTTCAAGGTCATGAT 120
 QY 121 CTCATGTTCCACTTCTCCACCTGCGGTGACATTTCTTGATGTCGTCCTCAT 180
 Db 121 CTCATGTTCCACTTCTCCACCTGCGGTGACATTTCTTGATGTCGTCCTCAT 180
 QY 181 CTGACCGAGGCCCATCAGACCTTTCGAGGACACCATCAAGGGCTTTGAGTGGCCCA 240
 Db 181 CTGACCGAGGCCCATCAGACCTTTCGAGGACACCATCAAGGGCTTTGAGTGGCCCA 240
 QY 241 CGAGAGTATCGGGTGTGTGATCCAGGGGATATATGTCCTCCCAAAATGCTACCTTAA 300
 Db 241 CGAGAGTATCGGGTGTGTGATCCAGGGGATATATGTCCTCCCAAAATGCTACCTTAA 300
 QY 301 TTATTTATCTTTAGATATTTATTTTGAAGAAAATACAAATCTTATCTTTGTGA 360
 Db 301 TTATTTATCTTTAGATATTTATTTTGAAGAAAATACAAATCTTATCTTTGTGA 360

QY 661 ATGTGTTTATATACATTTTCTCTCTCAATAGAGATTTTCTCCGATTTTATAA 720
DB 661 ATGTGTTTATATACATTTTCTCTCTCAATAGAGATTTTCTCCGATTTTATAA 720
QY 721 ATGACTATTAAGTCATTTTATTAAGACAGCATGTCGTAGATTCTCGTCAAAAATC 780
DB 721 ATGACTATTAAGTCATTTTATTAAGACAGCATGTCGTAGATTCTCGTCAAAAATC 780
QY 781 TTTTGATTTTTTAAAGCTAGTTGGCAACCCGTCTTCTCAAGAATTTGATTT 840
DB 781 TTTTGATTTTTTAAAGCTAGTTGGCAACCCGTCTTCTTCAAGAATTTGATTT 840
QY 841 TTCAAAAAAATAGTTATTTTCTCTTATTAATAAGAAAACATTAGAAAAATAGT 900
DB 841 TTCAAAAAAATAGTTATTTTCTCTTATTAATAAGAAAACATTAGAAAAATAGT 900
QY 901 TGGCAGACTAGCCCTTGAATGTTTCCCAATAAATTACATCATCTGTATATAATTG 960
DB 901 TGGCAGACTAGCCCTTGAATGTTTCCCAATAAATTACATCATCTGTATATAATTG 960
QY 961 GGCAGCCCATTAATTTTAAACGAAATCGAAGCGAAACCAATCTGAGCTAT 1020
DB 961 GGCAGCCCATTAATTTTAAACGAAATCGAAGCGAAACCAATCTGAGCTAT 1020
QY 1021 TTCTCTAGATTAGTAAAGAGAGAGAGAGAGAGAGAAATCAATTAAAGTATGTC 1080
DB 1021 TTCTCTAGATTAGTAAAGAGAGAGAGAGAGAGAGAAATCAATTAAAGTATGTC 1080
QY 1081 TGAAGATGTCGGTTTGGCAACGATAGCCACCGTATCATAGCTCATAGTGTCCATG 1140
DB 1081 TGAAGATGTCGGTTTGGCAACGATAGCCACCGTATCATAGCTCATAGTGTCCATG 1140
QY 1141 GGTGGGAGCTCTGTCATCTCATAGGAGCATATGATGTTGAAACGGTTCGTC 1200
DB 1141 GGTGGGAGCTCTGTCATCTCATAGGAGCATATGATGTTGAAACGGTTCGTC 1200
QY 1201 TTGTTCCATCTGTCAGACCTTGGCTATTTCTGAACCAAGAGATACCTATCCCAACAT 1260
DB 1201 TTGTTCCATCTGTCAGACCTTGGCTATTTCTGAACCAAGAGATACCTATCCCAACAT 1260
QY 1261 CCATCTTACTCATGCAACTTCCATGCAACAGCATATGTTTCTGTAAC 1311
DB 1261 CCATCTTACTCATGCAACTTCCATGCAACAGCATATGTTTCTGTAAC 1311

RESULT 4
BD062176 1394 bp DNA linear PAT 27-AUG-2002
LOCUS BD062176 Male tissue-preferred regulatory region and method of using same.
ACCESSION BD062176.1 GI:22607781
VERSION JP 2001520523-A/1
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1. (bases 1 to 1394) Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
AUTHORS Albertsen, M.C., Fox, T.W., Garnaat, C.W., Huffman, G.A. and Kendall, T.L.
TITLE Male tissue-preferred regulatory region and method of using same
JOURNAL PIONEER HI BRED INTERNATIONAL INC
COMMENT PN JP 2001520523-A/1
PD 30-OCT-2001
PF 19-JUN-1998 JP 1999504910
PR 23-JUN-1997 US 08/880499
PI MARC C ALBERTSEN, TIMOTHY W FOX, CARL W GARNAAAT, GARY A HUFFMAN, PI TIMMY L KENDALL
PC C12N15/82, C12N15/29, C12N9/24, C12N9/22, C12N9/10, C12N9/00 PC
PC C07K14/34, C12Q1/68,
PC A01H5/00
CC Strandedness: Single;
CC Topology: Linear;

FEATURES FH Key Location/Qualifiers.
source 1..1394
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 94.0%; Score 1311; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCATGTCCTCTATGAAAAAGATAGTACATGTCATATTCGTTTCTTAGAGTCC 60
DB 1 CCATGTCCTCTATGAAAAAGATAGTACATGTCATATTCGTTTCTTAGAGTCC 60
QY 61 CTTCTTGCCTTATTAAGCTGAATCGGGTTTCAAAAACTTCAAGGGTGCAATGAT 120
DB 61 CTTCTTGCCTTATTAAGCTGAATCGGGTTTCAAAAACTTCAAGGGTGCAATGAT 120
QY 121 CTCGATGTTCCACTTCTCCACCTCGGTGACATTTCTTGAGTGGTGGTCCCAT 180
DB 121 CTCGATGTTCCACTTCTCCACCTCGGTGACATTTCTTGAGTGGTGGTCCCAT 180
QY 181 CTGACCGAGGCCCATCAGACACCTTCCGGACACCCATCAAGGGCTTTCCGATGGCCA 240
DB 181 CTGACCGAGGCCCATCAGACACCTTCCGGACACCCATCAAGGGCTTTCCGATGGCCA 240
QY 241 CGAGACGTATCGGTCGTGTGTATCCAGGGATATATGTCCGCCAATCGTCACTATA 300
DB 241 CGAGACGTATCGGTCGTGTGTATCCAGGGATATATGTCCGCCAATCGTCACTATA 300
QY 301 TTATATCTTATGATATATATTTTATTTTGGAAAAATTAACAACTTATCTTTGTGTA 360
DB 301 TTATATCTTATGATATATATTTTATTTTGGAAAAATTAACAACTTATCTTTGTGTA 360
QY 361 GGGGCTCAGCATATGATTTTCCGTAGGGCCCAAGAAATCGAGAGACCAACCATGTCTAGT 420
DB 361 GGGGCTCAGCATATGATTTTCCGTAGGGCCCAAGAAATCGAGAGACCAACCATGTCTAGT 420
QY 421 TCCACTATTTGGCACTACCCAGAACAGATTTTAAAAAATTAACAAAGTAACTAATCCA 480
DB 421 TCCACTATTTGGCACTACCCAGAACAGATTTTAAAAAATTAACAAAGTAACTAATCCA 480
QY 481 CGAAGCTATCATGTAATGTTTAAAGAAACATCTTAAACCCAGATCCCTTAAAAAA 540
DB 481 CGAAGCTATCATGTAATGTTTAAAGAAACATCTTAAACCCAGATCCCTTAAAAAA 540
QY 541 CAAGCATATTTGAAAGAGACAAATTTATGTTTACAGTTTACAAACATCTTAAGCGACAAA 600
DB 541 CAAGCATATTTGAAAGAGACAAATTTATGTTTACAGTTTACAAACATCTTAAGCGACAAA 600
QY 601 TTATATCGAAAGTAAAGCTATGACGTTTCAATTTTCTTTTCAATCTTGTATTTTGT 660
DB 601 TTATATCGAAAGTAAAGCTATGACGTTTCAATTTTCTTTTCAATCTTGTATTTTGT 660
QY 661 ATGTGTTTATATACATTTTCTCTCTCAATAGAGATTTTCTCCGATTTTATAA 720
DB 661 ATGTGTTTATATACATTTTCTCTCTCAATAGAGATTTTCTCCGATTTTATAA 720
QY 721 ATGACTATTAAGTCATTTTATTAAGACAGCATGTCGTAGATTCTCGTCAAAAATC 780
DB 721 ATGACTATTAAGTCATTTTATTAAGACAGCATGTCGTAGATTCTCGTCAAAAATC 780
QY 781 TTTTGATTTTTTAAAGCTAGTTGGCAACCCGTCTTCTTCAAGAATTTGATTT 840
DB 781 TTTTGATTTTTTAAAGCTAGTTGGCAACCCGTCTTCTTCAAGAATTTGATTT 840
QY 841 TTCAAAAAAATAGTTATTTTCTCTTATTAATAAGAAAACATTAGAAAAATAGT 900
DB 841 TTCAAAAAAATAGTTATTTTCTCTTATTAATAAGAAAACATTAGAAAAATAGT 900
QY 901 TGGCAGACTAGCCCTTGAATGTTTCCCAATAAATTACATCATCTGTATATAATTG 960

Db 901 TGGCAGCTGACCTGAGAAATGTTTCCCAATTAATTCATCTGATGTTAATTTTGG 960
Qy 961 GCCAGCCCCATTAATTAATTTAAACCGAACTGAATGAGCGAAACCAATCTGAGCTAT 1020
Db 961 GCCAGCCCCATTAATTAATTTAAACCGAACTGAATGAGCGAAACCAATCTGAGCTAT 1020
Qy 1021 TTCTCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1080
Db 1021 TTCTCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1080
Qy 1081 TGAGATGCGGTTTGGCAACGATAGCCACCGTAATCATAGTATGAGTGGCTAGCTCA 1140
Db 1081 TGAGATGCGGTTTGGCAACGATAGCCACCGTAATCATAGTATGAGTGGCTAGCTCA 1140
Qy 1141 GGTTCGGCAGCTCTCGTGTCACTCATGSCATACACTACTGCTTGTCAACCGTTGCTC 1200
Db 1141 GGTTCGGCAGCTCTCGTGTCACTCATGSCATACACTACTGCTTGTCAACCGTTGCTC 1200
Qy 1201 TTGTTCCATGCTGCAAGCCTTGGCTATTCGAAACCAAGAGATACCTACTCCCAACAT 1260
Db 1201 TTGTTCCATGCTGCAAGCCTTGGCTATTCGAAACCAAGAGATACCTACTCCCAACAT 1260
Qy 1261 CCATCTTACTCATGCAACTTCATGCAACGACGACATATGTTCTCGAAC 1311
Db 1261 CCATCTTACTCATGCAACTTCATGCAACGACGACATATGTTCTCGAAC 1311

RESULT 5
LOCUS AF360356 3343 bp DNA linear PLN 12-MAY-2001
DEFINITION Zea mays male fertility protein (Me45) gene, complete cds.
ACCESSION AF360356
VERSION AF360356.1 GI:14028756
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 3343)
AUTHORS Fox,T.W., Trimnell,M.R. and Albertsen,M.C.
JOURNAL Unpublished
TITLE Cloning of Me45, a gene required for male fertility from Zea mays
REFERENCE 2 (bases 1 to 3343)
AUTHORS Fox,T.W., Trimnell,M.R. and Albertsen,M.C.
JOURNAL Direct Submission
TITLE Submitted (13-MAR-2001) Trait and Technology Development, Pioneer
JOURNAL HI-Bred Intl. Inc., 7300 N.W. 62nd Ave., P.O. Box 1004, Johnston,
IA 50131-1004, USA
FEATURES
source
Location/Qualifiers
1. .3343
/organism="Zea mays"
/mol_type="genomic DNA"
/cul_tivar="B73"
/db_xref="taxon:4577"
/chromosome="9L"
<1392..>2942
/gene="Me45"
join(<1392..1768,1898..2182,2280..2447,2534..>22942)
/gene="Me45"
/product="male fertility protein"
join(1392..1768,1898..2182,2280..2447,2534..2942)
/gene="Me45"
/codon_start=1
/product="male fertility protein"
/db_xref="GI:14028757"
/translation="MEKRLQMRGRGDIVQYPLHFAALALALVADPFGSLAEV
DYPVVKHLEAPGEVWMSPRDNASRLRRLRFVSGVFPBSIEFDLQKRGYAGIA
GAYVWVMEGAEETFAVNPVPMSEVCANGVNSTRKQHEKEBFGCGIAPFGE
TGEIYVADAYVGLVWVGSGVSAVARADSDPRLPANDLVDRNRSVGFPTDSRY
SRDHLNILLBEGTGRLALYDPTSGVNVVLKGLVFPNGVOISEDHQPLBSETTNC

ORIGIN

RIMRYVLEGPRAEVEVPAFLBGFPPNVSNRGOFWVAIDCCRTPAOEVPARPLR
TLYPKPBLKLVIMKARRMHTVIALLDGEGVVEVLDRGHEVMKLVSEVREVRK
LMIIVAHNHNIATIPYPLED"

Query Match 94.0%; Score 1311; DB 8; Length 3343;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CCATGCTGCTCTATGAAAAAGATGAGTCAATGTCTATATCCGTTTCTTGGGTC 60
Db 1 CCATGCTGCTCTATGAAAAAGATGAGTCAATGTCTATATCCGTTTCTTGGGTC 60
Qy 61 CTTCCTTCTGCTTATTAAGTCACTGAATCGGGGTTACAAAAAATTCCAGGGTCAAT 120
Db 61 CTTCCTTCTGCTTATTAAGTCACTGAATCGGGGTTACAAAAAATTCCAGGGTCAAT 120
Qy 121 CTCGATGTTCCACTTCTCCCACTCGCGGTGCAATTTCTGGATGCGGTGGTCCAT 180
Db 121 CTCGATGTTCCACTTCTCCCACTCGCGGTGCAATTTCTGGATGCGGTGGTCCAT 180
Qy 181 CTGACCGAGGCCCATCAGACACCTTTGGGAGACCCCATCAAGGCTTTGAGTGGCCA 240
Db 181 CTGACCGAGGCCCATCAGACACCTTTGGGAGACCCCATCAAGGCTTTGAGTGGCCA 240
Qy 241 CGAGACGTATCGGGTGTGATGATCGAGGATATATGTCCTCCCAATGCTCACTTAT 300
Db 241 CGAGACGTATCGGGTGTGATGATCGAGGATATATGTCCTCCCAATGCTCACTTAT 300
Qy 301 TTATTTATCTTATGATTAATTAATTTTGGAAAAATTAACAACCTTATCTTGTAT 360
Db 301 TTATTTATCTTATGATTAATTAATTTTGGAAAAATTAACAACCTTATCTTGTAT 360
Qy 361 GGGGCTCAGCATGATTTTCGCTTAGGGCCAGAAATGCGAGGACCGACATGCTAGTG 420
Db 361 GGGGCTCAGCATGATTTTCGCTTAGGGCCAGAAATGCGAGGACCGACATGCTAGTG 420
Qy 421 TCCACTATTGGCACTACCAAGAACAAATTAATTAATTAATTAATTAATTAATTAAT 480
Db 421 TCCACTATTGGCACTACCAAGAACAAATTAATTAATTAATTAATTAATTAATTAAT 480
Qy 481 CGAAACCTATCATGTAATGTTTAAAGAAACATCTATTAACGAGATCCTTAATAAAA 540
Db 481 CGAAACCTATCATGTAATGTTTAAAGAAACATCTATTAACGAGATCCTTAATAAAA 540
Qy 541 CAAGCATTTTGCAGAGAGACAAATTAATTAATTAATTAATTAATTAATTAATTAAT 600
Db 541 CAAGCATTTTGCAGAGAGACAAATTAATTAATTAATTAATTAATTAATTAATTAAT 600
Qy 601 TTATATCGAAAGGTAGCTATGACGTTCAAGATTTTCTTTTCATTTCTGTTATTTG 660
Db 601 TTATATCGAAAGGTAGCTATGACGTTCAAGATTTTCTTTTCATTTCTGTTATTTG 660
Qy 661 ATTGTTTATATACATTTTCTTCTTCAATAGTATTTCTTCCGATTTTATAA 720
Db 661 ATTGTTTATATACATTTTCTTCTTCAATAGTATTTCTTCCGATTTTATAA 720
Qy 721 ATGACTATTAAGTCAATTTTATATAGACAGCATGTCGATTTCTGTTCAAAAATC 780
Db 721 ATGACTATTAAGTCAATTTTATATAGACAGCATGTCGATTTCTGTTCAAAAATC 780
Qy 781 TTTCGATTTTATTAAGACCTAGTTTGGCAACCTGTTTCTTTCAAGAAATTTGATTT 840
Db 781 TTTCGATTTTATTAAGACCTAGTTTGGCAACCTGTTTCTTTCAAGAAATTTGATTT 840
Qy 841 TTCAAAAAAATTAATTAATTTTCTTCTTATTAATAAGAAACATTAATAAATAGAT 900
Db 841 TTCAAAAAAATTAATTAATTTTCTTCTTATTAATAAGAAACATTAATAAATAGAT 900
Qy 901 TGCAGACTAGCCCTGAGATGTTTCCCAATTAATTAATCAATGCTGTATTAATTTG 960
Db 901 TGCAGACTAGCCCTGAGATGTTTCCCAATTAATTAATCAATGCTGTATTAATTTG 960

OY	961	GCACACCCCTAAATTAATTTAAACCGAACTGAAATCGACGAAACCAATCTGAGCTAT	10202
Db	961	GCACACCCCTAAATTAATTTAAACCGAACTGAAATCGACGAAACCAAACTGAGCTAT	10202
OY	1021	TTCTCTAGATTAGTAAAGGAGAGAGAGAGAAATCAGTTTAAGTCATTGTGCC	10808
Db	1021	TTCTCTAGATTAGTAAAGGAGAGAGAGAGAAATCAGTTTAAGTCATTGTGCC	10808
OY	1081	TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTCTACGTCA	11404
Db	1081	TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTCTACGTCA	11404
OY	1141	GGTTGCGAGCTCTCGTGCATCTTCACATGCGATCATAGCTGTTTCAACCGTTGCTC	12000
Db	1141	GGTTGCGAGCTCTCGTGCATCTTCACATGCGATCATAGCTGTTTCAACCGTTGCTC	12000
OY	1201	TTGTTTCATCGTTCMAAGCTTGCCCTAATCTGAAACCAAGGATACCTACTCCCAACAT	12606
Db	1201	TTGTTTCATCGTTCMAAGCTTGCCCTAATCTGAAACCAAGGATACCTACTCCCAACAT	12606
OY	1261	CCATCTTACTCATGCACTTCATGCAACAGGACATATGTTTCCCTAAC	1311
Db	1261	CCATCTTACTCATGCACTTCATGCAACAGGACATATGTTTCCCTAAC	1311

RESULT	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
6	AX224402	AX224402 Sequence 9 from Patent WO0160997.	AX224402	AX224402.1	GI:15554644	Zea mays Zea mays	
		255 bp	DNA				linear PAT 10-SEP-2001

REFERENCE
AUTHORS
TITLE
JOURNAL

1. Albertsen, M.C., Fox, T.W., Garnaat, C.W., Huffman, G., and Kendall, T.L. Male tissue-preferred regulatory region and method of using same Patent: WO 060997-A 9 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)

Spinaetophyta; *Viridiplantae*; *Streptophyta*; *Embryophyta*; *Tracheophyta*; *Spermatophyta*; *Magnoliophyta*; *Liliopsida*; *Poales*; *Poaceae*; *PACCAD*
Clade; *Panicetidae*; *Andropogoneae*; *Zea*.

FEATURES	location/Qualifiers
source	1..255
	/Organism="Zea mays"
	/mol_type="unassigned DNA"
	/db_xref="taxon:4577"

ORIGIN					
Query Match	17.2%	Score 240;	DB 6;	Length 255;	
Best Local Similarity	100.0%	Pred. No. 2e-113;			
Matches 240; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;	

QY 1155 CGTGCATCTCACATGCGATACACTACATGCTTGTTCACCGCTTCGCTTGTTCATCGTCC 1214

Db 15 CGTGCATCTCACATGCGATACACTACATGCTTGTTCACCGCTTCGCTTGTTCATCGTCC 74

Oy		1215	AAGCCTTGCCATTCTGAAACCAAGAGTACTCTA	TCCCAACAATCATCTTA	CTATG	1274
D6		75	AAGCTTGCCATTCTGAAACCAAGAGTACTCTA	TCCCAACAATCATCTTA	CTATG	134

Oy 1275 CAACCTTCATGCAAAACAGCAGCATATGTTTCTGTAACAGATCTATTAAAGTCACACAG 1334
 Db 135 CAACCTTCATGCAAAACAGCAGCATATGTTTCTGTAACAGATCTATTAAAGTCACACAG 194

Oy 1335 CTAGGTTCTCCCGTAGCTTCCCTCTCCCTGCGCATCTTTTGGTCACCAACATG 1394
 |||||
 Db 195 CTAGCGTCTCCCGTAGCTTCCCTCTCCCTGCGCATCTTTTGGTCACCAACATG 254

RESULT 7		
AX224396		
LOCUS	158 bp	DNA
AX224396		linear
		PAT 10-SEP-2001

```

DEFINITION Sequence 3 from Patent WO0160997.
ACCESSION AX224396
VERSION AX224396.1 GI:15554638
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
PIONEER
FEATURES
SOURCE
1. .158
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

```

Query Match	100%	Score 111	DB 6	Length 158
Best Local Similarity	100.0%	Pred. No. 4	8e-46	
Matches 111	Conservative 0	Mismatches 0	Indels 0	Gaps 0

OY	1201	TTGTTTCATGTCGAACCGTTGCTATTTCGAAACCAAGAGATACCTACTCCCAACAAAT	1260
Db	48	TTGTTTCATGTCGCAACCGCTTGCTATTTCGAAACCAAGAGATACCTACTCCCAACAAAT	107
OY	1261	CCATCTTACTCATGCAACTTCATGCAAAACGACATATATGTTCTCGAAC	1311
Db	108	CCATCTTACTCATGCAACTTCATGCAAAACGACATATATGTTCTCGAAC	158

AX224398	50 bp	DNA	linear	PAT 10-SEP-2001
AX224398	Sequence 5	from Parent W00160997.		
AX224398				
AX224398.1	GI:15554640			
KEYWORDS				
SOURCE				
ORGANISM	Zea mays			
	Zea mays			

REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES

Eukaryotes: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

1
Albertsen, M. C., Fox, T. W., Garnaat, C. W., Huffman, G. and Kendall, T. L. Male tissue-preferred regulatory region and method of using same Patent: WO 0160997-A 5 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
Location/Qualifiers

```

source
1. .50
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

```

ORIGIN	Query Match	Similarity	3.6%;	Score 50;	DB 6;	Length 50;
	Best Local	100.0%;	Pred. No. 3.6e-14;			
	Matches	50;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
OY	1229	CTGACCAGAGGATCTACTCTCTCCCAACATCATCTTACTCATGCAAC				1278
db	1	CTGACCAGAGGATCTACTCTCTCCCAACATCATCTTACTCATGCAAC				50

LOCUS	Ax224339	40 bp	linear	FAI 10-SEP-2001
DEFINITION	Sequence 6 from Patent WO016097.			
ACCESSION	AX224339			
VERSION	AX224339.1	GI:15554641		

```

KEYWORDS
SOURCE      Zea mays
ORGANISM    Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
              clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
1
AUTHORS     Albertsen, M.C., Fox, T.W., Garnaat, C.W., Huffman, G. and Kendall, T.L.
TITLE       Male tissue-preferred regulatory region and method of using same
JOURNAL     Patent: WO 0160997-A 6 23-AUG-2001;
              PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
source      1..40
              /organism="Zea mays"
              /mol_type="unassigned DNA"
              /db_xref="taxon:4577"

ORIGIN
Query Match      2.9%; Score 40; DB 6; Length 40;
Best Local Similarity 100.0%; Pred. No. 6.1e-09;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1239 AGGATACCTACTCCCAACATCATCTTACTCATGCAAC 1278
        |||||
        1 AGGATACCTACTCCCAACATCATCTTACTCATGCAAC 40

Db
RESULT 10
LOCUS    AX224397                      30 bp      DNA      linear      PAT 10-SEP-2001
DEFINITION
Sequence 4 from Patent WO0160997.
ACCESSION AX224397
VERSION   AX224397.1 GI:15554639
KEYWORDS
SOURCE
ORGANISM Zea mays
              Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
              clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
1
AUTHORS     Albertsen, M.C., Fox, T.W., Garnaat, C.W., Huffman, G. and Kendall, T.L.
TITLE       Male tissue-preferred regulatory region and method of using same
JOURNAL     Patent: WO 0160997-A 4 23-AUG-2001;
              PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
source      1..30
              /organism="Zea mays"
              /mol_type="unassigned DNA"
              /db_xref="taxon:4577"

ORIGIN
Query Match      2.2%; Score 30; DB 6; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1179 CATGCTTGTCACCGTTGCTTGTTCCA 1208
        |||||
        1 CATGCTTGTCACCGTTGCTTGTTCCA 30

Db
RESULT 11
LOCUS    AC149836                      114817 bp   DNA      linear      HTG 22-JUN-2004
DEFINITION
Zea mays clone ZMMBMC0496117, *** SEQUENCING IN PROGRESS ***; 3
unordered pieces.
ACCESSION AC149836
VERSION   AC149836.1 GI:49035067
KEYWORDS HTG; HTGS_PHASE1; HTGS_FUZZTOP; HTGS_ACTIVEPIN.
SOURCE
ORGANISM Zea mays
              Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
              clade; Panicoideae; Andropogoneae; Zea.

```

```

REFERENCE
1 (bases 1 to 114817)
AUTHORS     Birren, B., Nusbaum, C., Lander, E., Butler, E., Wing, R., Bhatti, A.K.
              and Messing, J.
TITLE       Zea mays, clone ZMMBMC0496117
JOURNAL     Unpublished
REFERENCE
2 (bases 1 to 114817)
AUTHORS     Birren, B., Nusbaum, C., Lander, E., Butler, E., Wing, R., Bhatti, A.K.,
              Messing, J., Aboueleil, A., Allen, N., Anderson, M., Anderson, S.,
              Archach, H.M., Barna, N., Baetien, V., Bloom, T., Boguslavsky, L.,
              Bouknight, B., Camarata, J., Chang, J., Choepel, Y., Collimore, A.,
              Cook, A., Cooke, P., Corum, B., Dearlano, K., Diaz, J.S., Dodge, S.,
              Doolley, K., Dorris, L., Erickson, J., Fero, S., Ferreira, P.,
              Fitzgerald, M., Gage, D., Galagan, J., Gardna, S., Graham, L.,
              Grand-Pierre, N., Haez, N., Hagopian, D., Hage, B., Hall, J.,
              Horton, A., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,
              Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,
              Liu, G., Liu, X., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P.,
              Major, J., Manning, J., Matthews, C., McCarthy, M., Meidrim, J.,
              Meneses, L., Minova, T., Miengo, V., Murphy, T., Naylor, J., Nguyen, C.,
              Nguyen, T., Nicol, R., Nordu, C., O'Connor, T., O'Donnell, P.,
              O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
              Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
              Roman, J., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C.,
              Spencer, B., Strange-Thomann, N., Stojanovic, N., Stubbs, M.,
              Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M.,
              Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
              Vyskocil, H., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
              Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
              Direct Submission
              Submitted (22-JUN-2004) Whitehead Institute/MIT Center for Genome
              Research, 320 Charles Street, Cambridge, MA 02141, USA
              All repeats were identified using RepeatMasker:
              Smit, A.F.A. & Green, P. (1996-1997)
              http://ftp.genome.washington.edu/RN/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/MIT Center for Genome Research
Center code: MIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@road.mit.edu

-----
Bhatti, AK and Messing, J: The Plant Genome Initiative at
Rutgers, Waksman Institute, Rutgers, The State University of New
Jersey, 190 Frelinghuysen Road, Piscataway, NJ 08854, USA
(http://pgir.rutgers.edu)
Butler, E and Wing, R: Arizona Genomics Institute, Biological
Sciences West, 448A, P.O. Box 210088, University of Arizona,
Tucson, AZ 85721, USA (http://www.genome.arizona.edu)
----- Project Information
Center project name: L10352
Center clone name: 496_L_17

-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1      35156: contig of 35156 bp in length
*      35157      35256: gap of 100 bp
*      35257      97940: contig of 62684 bp in length
*      97941      98040: gap of 100 bp
*      98041      114817: contig of 16777 bp in length.

Location/Qualifiers
1..114817
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ZMMBMC0496117"
/clone_lib="CHORI-201 Maize B73 BAC MDO1 Library"

ORIGIN

```

Query Match 1.8%; Score 25; Ds 2; Length 114817;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 844 AAAAAAAAAATTGTTATTTCTCTT 868
27886 AAAAAAAAAATTGTTATTTCTCTT 27910

RESULT 12
AE010607 10384 bp DNA linear BCT 25-MAR-2002
LOCUS Fusobacterium nucleatum subsp. nucleatum ATCC 25586, section 149 of
DEFINITION 197 of the complete genome.
ACCESSION AE010607 AE009851
VERSION AE010607.1 GI:19714578
KEYWORDS

SOURCE
ORGANISM
Fusobacterium nucleatum subsp. nucleatum ATCC 25586
Bacteria; Fusobacteriales; Fusobacteriaceae;
Fusobacterium.

REFERENCE
AUTHORS
1 (bases 1 to 10384)
Kaparat, V., Anderson, I., Ivanova, N., Reznik, G., Los, T.,
Lykidis, A., Bhattacharyya, A., Bartman, A., Gardner, W., Grechkin, G.,
Zhu, L., Vasileva, O., Chu, L., Kogan, Y., Chaga, O., Goldsman, E.,
Bernal, A., Larsen, N., D'Souza, M., Walunas, T., Pusch, G.,
Haselkorn, R., Fongstein, M., Kyriades, N. and Overbeek, R.
Genome sequence and analysis of the oral bacterium Fusobacterium
nucleatum strain ATCC 25586
J. Bacteriol. 184 (7), 2005-2018 (2002)

JOURNAL
MEDLINE
2186394
11889109
2 (bases 1 to 10384)
Kaparat, V., Anderson, I., Ivanova, N., Reznik, G., Los, T.,
Lykidis, A., Bhattacharyya, A., Bartman, A., Gardner, W., Grechkin, G.,
Zhu, L., Chu, L., Kogan, Y., Chaga, O., Goldsman, E., Bernal, A.,
Larsen, N., D'Souza, M., Walunas, T., Pusch, G.D., Haselkorn, R.,
Fongstein, M., Kyriades, N. and Overbeek, R.
Direct Submision
Submitted (13-FEB-2002) Integrated Genomics, 2201 W. Campbell Park
Drive, Chicago, IL 60612, USA
Location/Qualifiers

FEATURES
source
1. 10384
/organism="Fusobacterium nucleatum subsp. nucleatum ATCC
25586"

gene
CDS
162..1664
/gene="FN0998"
162..1664
/gene="FN0998"
/codon_start=1
/transl_table=11
/product="Dipeptide-binding protein"

gene
CDS
/db_xref="GI:19714579"
/protein_id="AAJ95194.1"
/db_xref="GI:19714579"
/translation="MKRTFLVLSIFLVGCGQNKTPSPNGSTVVIIGGARPKSLDP
HMYNSIPDLVSRQFNLTFRREKDSIRPELASVEYKNDKELDIILKGLKFDGS
ELTADVVPSFERMKDPGSSIVMEIIDVKEIKNDYEIKLFLKSSSALLNLAHPIIT
STLVKKYVAGNDLSVAPMGTAFLIANDGEKIELAFKQYPSGAPVKEKITPRAT
PBDTSMALALEGENDIATGMPVSTQTIENDKDLISEPTATATEYICLANEKVFPD
NKDFKALNYAIDKOSIIDSISRGKVAKSVNPFVFGYDGFSEYFPNPKAKBLI
KSGGVDMFSLYVNDSPVRLQVAQIIQANLQVGDANIETLEWGTLYQKTEGDFPS
AYLGGWISGTSADIVLYPLDLSKISGFGNARARYNPEDFGEVEMARIALPBERKE
HFKNQAIISOEDSPLIVLYKNKENIGINKRIKGFEDPTTMMKFKNLKLIK"

gene
CDS
1669..2712
/gene="FN0999"
1669..2712
/gene="FN0999"
/EC_number="3.4.11.-"

/codon_start=1
/transl_table=11
/product="Dehydrolocking aminopeptidase"
/protein_id="AAJ95195.1"
/db_xref="GI:19714580"

gene
CDS
2854..3936
/gene="FN1000"
/EC_number="2.8.1.6"
/codon_start=1
/transl_table=11
/product="Biotin synthase"
/protein_id="AAJ95196.1"
/db_xref="GI:19714581"

gene
CDS
3926..4585
/gene="FN1001"
/EC_number="6.3.3.3"
/codon_start=1
/transl_table=11
/product="Dehydrolocking synthetase"
/protein_id="AAJ95197.1"
/db_xref="GI:19714582"

gene
CDS
4653..6011
/gene="FN1002"
4653..6011
/gene="FN1002"
/EC_number="2.6.1.62"
/codon_start=1
/transl_table=11
/product="Adenosylmethionine-8-amino-7-oxononanoate
aminotransferase"

gene
CDS
6737..7558
/gene="FN1003"
6737..7558
/gene="FN1003"
/codon_start=1
/transl_table=11
/product="Outer membrane protein P1 precursor"
/protein_id="AAJ95199.1"
/db_xref="GI:19714584"

gene
CDS
6737..7558
/gene="FN1003"
6737..7558
/gene="FN1003"
/codon_start=1
/transl_table=11
/product="Outer membrane protein P1 precursor"
/protein_id="AAJ95199.1"
/db_xref="GI:19714584"

gene
CDS
6737..7558
/gene="FN1003"
6737..7558
/gene="FN1003"
/codon_start=1
/transl_table=11
/product="Outer membrane protein P1 precursor"
/protein_id="AAJ95199.1"
/db_xref="GI:19714584"

```

/translation="MDAATGKGLSAAOIAIAIKTKTNEALAKLQORVNLSONGSLG
DIOSKREMGEPOLGIVNYKNDKLANAARDSDIKANPAKAGHEHLETTDILKQIT
GLSTFYQYITNSKIRDLPAITLSVGASYKADNYLVSTTANIFNIOADMDRPTTIG
EHEGRDYKNGWEIAVENEYKLNKFTLIGSLNANGATASYNDEYALNSFTLGG
GIRYODESIALITASVAHFYQAGNFKKEYGTENQKYKKEITAVAGLSVTKEF"
7577. .8143
/gene="FN1004"
7577. .8143
/gene="FN1004"
/codon_start=1
/translation="Transcriptional regulator, Tetr family"
/product="Transcriptional regulator, Tetr family"
/protein_id="AA195201.1"
/db_xref="GI:19714585"
/translation="MPKKVLSKSFILDKSEFELKEGIGSISARNYKILDSAPDI
YKISGSMNKKEIKAKADLEFKKGTGIFLDIGKIGISIFABEQLPQVDS
KDNIEGLIDELFLIRREIKKORLIRIKKEQEBLLVCWFPAHGLSTLIATGFK
NPDEFEKYLADAPARLFYEIENYSK"
8337. .8810
/gene="FN1005"
8337. .8810
/gene="FN1005"
/codon_start=1
/translation="Hypothetical protein"
/product="Hypothetical protein"
/protein_id="AA195201.1"
/db_xref="GI:19714586"
/translation="MINQNAORFLDSOLOQLEVDORRERERAAAAQYQEPDRDI
YHRAVAVNMEDTGHFYGPPELLMTPTPNSSKAKAIKAKKEPRIGADLDIKATYH
LIMDGLRYIVMGTGNTKTKGMEAFKFEVDENIKILINKCSITONCDPFSNSN"
complement(8863. .9261)
/gene="FN1006"
complement(8863. .9261)
/gene="FN1006"
/EC_number="2.3.1.-"
/codon_start=1
/translation="Acetyltransferase"
/product="Acetyltransferase"
/protein_id="AA195202.1"
/db_xref="GI:19714587"
/translation="MECKIINKNTNYNLNDLTKLNTSYAKDRKKEKYKVEKSLC
YPAVDKNTKLGPARAITDTYNTYLCDIIVDEBYNGKIGKKLVETLLINDDLQYV
RGLITDKAKKPEKGFVNKEDVMQDKK"
9433. .10317
/gene="FN1007"
9433. .10317
/gene="FN1007"
/codon_start=1
/translation="unknown"
/product="unknown"
/protein_id="AA195203.1"
/db_xref="GI:19714588"
/translation="MKENNNNRKHKRDLVFLIWNILIFGLSIFLPSLANKKVDN
DIEVYFLIIGFILISLITIKIKIKAKAPNSKNITKTEKEHAYKPLNKKVDS
LVYRMVIVAPILFSFYIPLXIASKIQNRENEVSLIYGGKTFPIITLPCIV
IFPLIYSSQKIRMLQKVDFAKSEVNNLJDESLTELTPKXITPFRKPFIMWD
GSLNIFLIEDIEVEYKKNYFPIYTKLILHKNGRKKIRYAGPDENMERRKNPIV
KNSNIEGKEYHINLPI"

```

```

ORIGIN
Query Match 1.7% Score 24; DB 1; Length 10384;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 872 AAAATAGAAAACACTTAGAAAAAT 895
Db 2719 AAAATAGAAAACACTTAGAAAAAT 2742

```

```

RESULT 13 AC113237 184509 bp DNA 1linear HTG 07-AUG-2002
LOCUS AC113237/c
DEFINITION Canis familiaris clone RP01-70J12, WORKING DRAFT SEQUENCE, 3

```

```

ORDERED PIECES.
AC113237.2 GI:22128672
HTG; HTGS PHASE2; HTGS DRAFT.
KEYWORDS
SOURCE
ORGANISM
Canis familiaris
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euteria; Carnivora; Fissipedia; Canidae; Canis.
1 (bases 1 to 184509)
Ahter,N., Antonellis,A., Ayle,K., Beckstrom-Sternberg,S.M.,
Benjamin,B., Blakeley,R.W., Bouffard,G.G., Breen,K., Brinkley,C.,
Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J.,
Haghighi,P., Hansen,N., Ho,S.-L., Ido,J.R., Karlins,E., Laric,P.,
Lee-Jin,S.-O., Legaapi,R., Maduro,O.L., Maduro,V.B.,
Margulies,E.H., Mastello,C., Maekert,B., Mastrion,S.D.,
McCluskey,J.C., McDowell,J., Paguitigan,C., Pearson,R.,
Portnoy,M.E., Prasad,A., Schueler,M.G., Stantioop,S., Thomas,J.W.,
Thomas,P.J., Touchman,J.W., Tsurgon,C., Vogt,J.L., Walker,M.A.,
Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 184509)
Green,E.D.
Direct Submission
Submitted (27-FEB-2002) NIH Intramural Sequencing Center, 8717
Groveomont Circle, Galtersburg, MD 20877, USA
3 (bases 1 to 184509)
Green,E.D.
Direct Submission
Submitted (07-AUG-2002) NIH Intramural Sequencing Center, 8717
Groveomont Circle, Galtersburg, MD 20877, USA
On Aug 7, 2002 this sequence version replaced gi:18958664.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc_zoo@nigtl.nih.gov
----- Project Information
Center project name: avx
Center clone name: 070J12

```

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phred-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 184204 bases at least Q40
 Consensus quality: 184261 bases at least Q30
 Consensus quality: 184294 bases at least Q20
 Insert size: 16200; agarose-fp
 Insert size: 184309; sum-of-contigs
 Quality coverage: 10.51x in Q20 bases; agarose-fp
 Quality coverage: 9.24x in Q20 bases; sum-of-contigs

----- NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced
 * by the finished sequence as soon as it is available and


```

* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
  1 212621: contig of 212621 bp in length.
FEATURES
  source
    1..212621
      /organism="Rattus norvegicus"
      /mol_type="genomic DNA"
      /db_xref="taxon:10116"
      /clone="CH230-176E2"
      1828..2550
      /note="clone_boundary
      clone_end:T7
      site:ECORI
      end_sequence:BH359530"
  misc_feature
    end_sequence:BH359530"
ORIGIN
Query Match      1.7%  Score 24;  DB 2;  Length 212621;
Best Local Similarity 100.0%;  Pred. NO. 1.1;
Matches 24;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
QY      848 AAAATTAGTTATTTCTCTTAT 871
Db      1464 AAAATTAGTTATTTCTCTTAT 1441
RESULT 15
AC110715/c
LOCUS
DEFINITION
  Rattus norvegicus clone CH230-155J23, *** SEQUENCING IN PROGRESS
  *** 2 unordered pieces.
AC110715
AC110715.5 GI:25072769
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
  1 (bases 1 to 213753)
  Muzny,D,Marie, Metzker,M, Lee, Abramson,S, Adams,C, Alder,J,
  Allen,C, Allen,H, Alibrooke,S, Amin,A, Anguiano,D,
  Anyalebechi,V, Aoyagi,A, Ayodeji,M, Baca,E, Baden,H,
  Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benahmed,F,
  Blawie,K, Blair,J, Blankenburg,K, Blyth,P, Brown,M,
  Bryant,N, Buhay,C, Burch,P, Burrell,K, Calderon,E,
  Cardenas,V, Carter,K, Cavazos,I, Cessari,H, Center,A,
  Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chen,Z, Chu,J,
  Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,L,
  Davila,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dedrich,D,
  Delgado,O, Denson,S, Deramo,C, Ding,Y, Dinh,C, Diya,K,
  Draper,H, Dugan-Rocha,S, Dunn,A, Durbin,K, Duval,B, Eaves,K,
  Egan,A, Escotto,M, Eugene,C, Evans,C,A, Fella,T, Fan,G,
  Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,M, Foster,P,
  Frazer,C,M, Gabler,A, Ganta,R, Garcia,A, Garner,T, Garza,M,
  Gebregeorgis,E, Geer,K, Gill,R, Grady,M, Guerra,W, Guevara,M,
  Gunaratne,P, Haaland,W, Hamill,C, Hamilton,C, Hamilton,K,
  Harvey,Y, Havlak,P, Hawes,A, Henderson,N, Hernandez,J,
  Hernandez,R, Hines,S, Hladun,S,L, Hodgson,A, Hoguee,M,
  Hollins,B, Howells,S, Huliy,S, Hume,J, Idlebird,D, Jackson,A,
  Jackson,L, Jacob,L, Jiang,H, Johnson,B, Johnson,R, Jolyvec,A,
  Karpathy,S, Kelly,S, Kelly,S, Khan,Z, King,L, Kovar,C,
  Kowib,C, Kraft,C,L, Ledow,H, Levan,J, Lewis,L, Li,Z, Liu,J,
  Liu,J, Liu,M, Liu,Y, London,P, Longacre,S, Lopez,D,
  Lorensbuewa,L, Louisedge,H, Lozada,R,J, Lu,X, Ma,J,
  Maheshwari,M, Mahindaratne,M, Mahmoud,M, Malloy,K, Mangum,A,

```

```

Mangum,B, Mapua,P, Martin,K, Martin,R, Martinez,E,
Mawhinney,S, McLeod,M,P, McNeill,T,Z, Meenen,E,
Miloasavilevich,A, Miner,G, Minga,E, Montemayor,J, Moore,S,
Morgan,M, Morris,K, Morris,S, Munidasa,M, Murphy,M, Nair,L,
Nankervis,C, Neal,D, Newton,N, Nguyen,N, Norris,S,
Nwachilemech,O, Okwumodu,G, Olampungsoon,A, Pal,S, Parks,K,
Pasternak,S, Paul,H, Perez,A, Perez,L, Frankoch,C,
Plopper,F, Poindexter,A, Popovic,D, Primus,E, Pu,L, L,
Piazo,M, Quiroz,J, Rachlin,E, Reeves,K, Regier,M,A, Reigh,R,
Reilly,B, Reilly,M, Ren,Y, Reuter,S, Richards,S, Riggs,F,
Rives,C, Rodkey,T, Rojas,A, Rose,M, Rose,R, Ruiz,S,J,
Sanders,M, Saverly,G, Scherer,S, Scott,G, Shltsman,S, Shen,H,
Sherry,J, Shvartsbeyn,A, Siseon,I, Sitter,C,D, Smajs,D,
Sneed,A, Sodergren,E, Song,X,-Z, Sorelle,K, Soea,U,
Steinle,M, Strong,R, Sutton,A, Svatek,A, Taber,P, Taylor,C,
Taylor,T, Thomas,N, Thomas,S, Tingey,A, Trejos,Z, Umanil,K,
Valas,R, Vera,V, Villaseana,D, Waldron,L, Walker,B, Wang,U,
Wang,O, Wang,S, Warren,J, Warren,R, Wei,X, White,F,
Williams,G, Willison,R, Wiczek,R, Woden,H, Worley,K,
Wright,D, Wright,R, Wu,J, Yakub,S, Yen,U, Yoon,L, Yoon,V,
Yu,F, Zhang,J, Zhou,J, Zhou,X, Zhao,S, Dunn,D, von
Niederhausern,A, Weiss,R, Smith,D,R, Holt,R,A, Smith,H,O,
Weinstock,G, and Gibbs,R,A.
  Direct Submission
Unpublished
  2 (bases 1 to 213753)
  Worley,K,C.
  Direct Submission
Submitted (15-FEB-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
  3 (bases 1 to 213753)
  Rat Genome Sequencing Consortium.
  Submitted (19-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
  On Nov 19, 2002 this sequence version replaced gi:23680201.
  The sequence in this assembly is a combination of BAC based reads
  and whole genome shotgun sequencing reads assembled using Atlas
  (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
  in the feature table below represents a scaffold in the Atlas
  assembly (a 'contig-scaffold'). Within each contig-scaffold,
  individual sequence contigs are ordered and oriented, and separated
  by sized gaps filled with Ns to the estimated size. The sequence
  may extend beyond the ends of the clone and there may be sequence
  contigs within a contig-scaffold that consist entirely of whole
  genome shotgun sequence reads. Both end sequences and whole genome
  shotgun sequence only contigs will be indicated in the feature
  table.
  ----- Genome Center
  Center: Baylor College of Medicine
  Center code: BCM
  Web site: http://www.hgsc.bcm.tmc.edu/
  Contact: hgsc-help@bcm.tmc.edu
  ----- Project Information
  Center project name: GSTRM
  Center clone name: CH230-155J23
  ----- Summary Statistics
  Assembly program: Phrap, version 0.990329
  Consensus quality: 203043 bases at least Q40
  Consensus quality: 204685 bases at least Q30
  Consensus quality: 205672 bases at least Q20
  Estimated insert size: 206221; sum-of-contigs estimation
  Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
  -----
  * NOTE: Estimated insert size may differ from sequence length
  * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
  * NOTE: This is a 'working draft' sequence. It currently
  * consists of 2 contigs. The true order of the pieces
  * is not known and their order in this sequence record is
  * arbitrary. Gaps between the contigs are represented as
  * runs of N, but the exact sizes of the gaps are unknown.

```


* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved

* 1 212625: contig of 212625 bp in length
* 212626 212725: gap of unknown length
* 212726 213753: contig of 1028 bp in length.
Location/Qualifiers

source 1. 213753
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-155J23"
1. 1535
misc_feature /note="wgs_end_extension
clone_end:Sp6"
1586. .2703
misc_feature /note="wgs_end_extension
clone_end:Sp6"
5028. .5742
misc_feature /note="clone boundary
clone_end:Sp6
site:
end_sequence: BH364626"
33891. .35390
misc_feature /note="wgs_contig"
21446. .212625
misc_feature /note="wgs_contig"

ORIGIN

Query Match 1.7%; Score 24; DB 2; Length 213753;
Best Local Similarity 100.0%; Pred.No. 1.1;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 848 AAAATTAGTTATTCTCTTTAT 871
|||||
DB 65390 AAAATTAGTTATTCTCTTTAT 65367

Search completed: September 16, 2005, 15:25:40
Job time : 4107 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2005, 08:08:50 ; Search time 578.5 Seconds
(without alignments)
14264.684 Million cell updates/sec

Title: US-10-713-381-2

Perfect score: 1394
Sequence: 1 cccctggcgtcctcatgaaaa.....cttttcgcccaccaccatg 1394

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

N_Geneseq_16Dec04:*
1: geneeqn19808:*
2: geneeqn19908:*
3: geneeqn20008:*
4: geneeqn20018:*
5: geneeqn20018:*
6: geneeqn20028:*
7: geneeqn20028:*
8: geneeqn20038:*
9: geneeqn20038:*
10: geneeqn20038:*
11: geneeqn20038:*
12: geneeqn20048:*
13: geneeqn20048:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1394	100.0	1394	2	AAx07409
2	1394	100.0	1394	2	AAH76333
3	1311	94.0	1394	2	AAx07408
4	1311	94.0	1394	2	AAH76332
5	240	17.2	255	5	AAH76340
6	111	8.0	158	5	AAH76334
7	50	3.6	50	5	AAH76336
8	40	2.9	40	5	AAH76337
9	30	2.2	30	5	AAH76335
10	23	1.6	849	6	ABL40437
11	21	1.5	333	5	ABV56505
12	21	1.5	459	4	ABA58867
13	21	1.5	459	4	AAI38581
14	21	1.5	459	4	AAK32768
15	21	1.5	459	4	AAK07027
16	21	1.5	459	4	ABs32487
17	21	1.5	459	6	ABs07566
18	21	1.5	965	3	AAa59660
19	21	1.5	9255	6	ABL34152
20	21	1.5	9265	6	AAa53354

21	21	1.5	18817	6	ABL34494	ABL34494 Human met
22	21	1.5	18817	6	ABL70161	ABL70161 Chemical
23	21	1.5	18817	7	ADs99755	ADs99755 Bisulphite
24	21	1.5	26493	12	ADJ12386	ADJ12386 DNA fragm
25	20	1.4	260	5	ABV60726	ABV60726 Human pro
26	20	1.4	680	6	ABO18134	ABO18134 Oligonuc
27	20	1.4	680	6	ABO18135	ABO18135 Oligonuc
28	20	1.4	2523	6	AAH76337	AAH76337 Zea may
29	20	1.4	10439	6	AAH7817	AAH7817 Alternati
30	20	1.4	11779	6	AAH42891	AAH42891 Maize sug
31	20	1.4	51198	11	ACN45000	ACN45000 Mouse gen
32	20	1.4	91071	11	ACN44004	ACN44004 Mouse gen
33	20	1.4	110000	6	ABA90193_2	Continuation (3 of
34	20	1.4	110000	6	ABA90193_3	Continuation (4 of
35	20	1.4	110000	6	ABQ87681_2	Continuation (3 of
36	20	1.4	110000	6	ABQ87681_3	Continuation (4 of
37	20	1.4	110000	8	ABX33717_2	Continuation (3 of
38	20	1.4	110000	8	ABX33717_3	Continuation (4 of
39	20	1.4	148497	12	ADO18925	Add18925 Human bof
40	20	1.4	194534	12	ADO97481	ADO97481 Human can
41	20	1.4	337344	13	ABD32715	ABD32715 Human can
42	19	1.4	418	13	ADR65258	ADR65258 Cotton CD
43	19	1.4	468	2	AAH19065	AAH19065 Human PPA
44	19	1.4	485	4	ABA57528	ABA57528 Human PPA
45	19	1.4	485	4	AAI37083	AAI37083 Human foe

ALIGNMENTS

RESULT 1	AAx07409	standard; DNA; 1394 BP.
ID	AAx07409	standard; DNA; 1394 BP.
AC	AAx07409;	
XX		
DT	08-JUN-1999	(first entry)
XX		
DE	Zea may	
XX		
DE	Zea may	Me45 male tissue-preferred regulatory region.
XX		
KW	Me45; male; tissue-preferred; regulatory region; plant cells;	
KM	plant tissue; differentiated; hybrid seed; fertility; ss.	
XX		
OS	Zea may.	
XX		
PN	W09859061-A1.	
XX		
PD	30-DEC-1998.	
XX		
PF	19-JUN-1998;	98WO-US012895.
XX		
PR	23-JUN-1997;	97US-00880499.
XX		
PA	(PION-) PIONEER HI-BRED INT INC.	
XX		
PI	Albertsen MC, Fox TW, Garmaat CW, Huffman GA, Kendall TL;	
XX		
DR	WPI; 1999-105628/09.	
XX		
PT	New nucleic acid encoding a Me45 male tissue-preferred regulatory region	
XX	- useful in mediating plant fertility, especially hybrid seed production.	
XX		
PS	Claim 3; Page 23-24; 39pp; English.	
XX		
CC	The sequence is that encoding an Me45 male tissue-preferred regulatory	
CC	region. It may be used in the construction of a vector for a method of	
CC	producing exogenous genes in a male tissue-preferred manner, which is	
CC	useful in restoring or conferring fertility, such as in hybrid seed	
CC	production. In conferring fertility, a monocot/dicot plant is transformed	
CC	with the exogenous nucleotide sequence (a male sterility gene, preferably	
CC	Me45), which encodes a product selected from auxins, rols and diptheria	
CC	toxin. Hybrid seeds are produced by cross-pollinating maize male fertile	
CC	and infertile plants	

XX Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other:
 Query Match 100.0%; Score 1394; DB 2; Length 1394;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCATGGGTCTCTATGAAAAAGATGAGTACATATGTCTATATCCGTTTTCTTAAGGATCC 60
 DB 1 CCAATGGTCTCTATGAAAAAGATGAGTACATATGTCTATATCCGTTTTCTTAAGGATCC 60
 QY 61 CTTCTTCTGCTTATTAATGATCGAATCGGGGTTACAAAAACCTTCCACGGGTGATAT 120
 DB 61 CTTCTTCTGCTTATTAATGATCGAATCGGGGTTACAAAAACCTTCCACGGGTGATAT 120
 QY 121 CTCGATGTTCCACTTCTCCCACTCGGGTTCACATTTCTTGGATGTGGGTGTTCCCAT 180
 DB 121 CTCGATGTTCCACTTCTCCCACTCGGGTTCACATTTCTTGGATGTGGGTGTTCCCAT 180
 QY 181 CTGACCGAGGCCCATCAGACACCTTTCGGGACCCCATCAAGGGCCCTTTCGATGGCCCA 240
 DB 181 CTGACCGAGGCCCATCAGACACCTTTCGGGACCCCATCAAGGGCCCTTTCGATGGCCCA 240
 QY 241 CGAGACGTATCGGGTGTGTGATCCAGGGGATATATGTCCTCCACATGTCACCTATA 300
 DB 241 CGAGACGTATCGGGTGTGTGATCCAGGGGATATATGTCCTCCACATGTCACCTATA 300
 QY 301 TTATTTATCTTTAGATATTTATTTATTTTGGAAAAATAACAATTATATCTTTGTGTA 360
 DB 301 TTATTTATCTTTAGATATTTATTTATTTTGGAAAAATAACAATTATATCTTTGTGTA 360
 QY 361 GGGCCTCAGCATAGATTTTGGTGAAGGCCCCAGAAAAATGCGAGACCGCATGTAGTG 420
 DB 361 GGGCCTCAGCATAGATTTTGGTGAAGGCCCCAGAAAAATGCGAGACCGCATGTAGTG 420
 QY 421 TCCACTATTGGCACTACCCAGAACAGATTTAAAAAATACCAAGTAACTAATCCACT 480
 DB 421 TCCACTATTGGCACTACCCAGAACAGATTTAAAAAATACCAAGTAACTAATCCACT 480
 QY 481 CGAAAGCTATCAGTATATGTTTTAAAGAAAAACATATTAACACGATCCTCTAATAAAA 540
 DB 481 CGAAAGCTATCAGTATATGTTTTAAAGAAAAACATATTAACACGATCCTCTAATAAAA 540
 QY 541 CAAGCATATTTGCAAAAGAGACAAATATGTATACAGTTTCAAAACATCTAAGACGCAAA 600
 DB 541 CAAGCATATTTGCAAAAGAGACAAATATGTATACAGTTTCAAAACATCTAAGACGCAAA 600
 QY 601 TTATATCGAAAGGTAAAGTATGACGTTTCAATTTCTTTTCAATCTTGTATATTTGTT 660
 DB 601 TTATATCGAAAGGTAAAGTATGACGTTTCAATTTCTTTTCAATCTTGTATATTTGTT 660
 QY 661 ATTTGTTTTTATATACATTTTCTCTTCAATAGAGTATTTTCTTCCGATTTTATAA 720
 DB 661 ATTTGTTTTTATATACATTTTCTCTTCAATAGAGTATTTTCTTCCGATTTTATAA 720
 QY 721 ATGACTATTAAGTATTTTATATTAAGAGCAGCATGTGTATCTCGTCAAAAATC 780
 DB 721 ATGACTATTAAGTATTTTATATTAAGAGCAGCATGTGTATCTCGTCAAAAATC 780
 QY 781 TTTTGTATTTTAAAGCTAGTTGGCAACCTGTTTCTTCAAGAAATTTGATTTT 840
 DB 781 TTTTGTATTTTAAAGCTAGTTGGCAACCTGTTTCTTCAAGAAATTTGATTTT 840
 QY 841 TTCAAAAAAATAGTTTATTTCTCTTATTAATAAGAAAAACATTAAGAAAAATGAGT 900
 DB 841 TTCAAAAAAATAGTTTATTTCTCTTATTAATAAGAAAAACATTAAGAAAAATGAGT 900
 QY 901 TGGCAGACTAGCCCTAGAAATGTTTCCCATATAATTAACATCACTGTATATAATTTTG 960
 DB 901 TGGCAGACTAGCCCTAGAAATGTTTCCCATATAATTAACATCACTGTATATAATTTTG 960
 QY 961 GCCAGCCCCATTAATTTTAAACCGAAACTGAATCGAGCAACCAATCTGACCTAT 1020
 DB 961 GCCAGCCCCATTAATTTTAAACCGAAACTGAATCGAGCAACCAATCTGACCTAT 1020

DB 961 GCCAGCCCCATTAATTTTAAACCGAAACTGAATCGAGCAACCAATCTGACCTAT 1020
 QY 1021 TTCTCTAGATTTGTAATAAAGGAGAGAGAGAAATCATGTTTAAGTATGTCCC 1080
 DB 1021 TTCTCTAGATTTGTAATAAAGGAGAGAGAGAAATCATGTTTAAGTATGTCCC 1080
 QY 1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCTAGTCA 1140
 DB 1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCTAGTCA 1140
 QY 1141 GGTTCGGACGCTCTCGTGCATCTCATAGGATATACATGCTTGTTCMAACGTTCTGTC 1200
 DB 1141 GGTTCGGACGCTCTCGTGCATCTCATAGGATATACATGCTTGTTCMAACGTTCTGTC 1200
 QY 1201 TTGTTCCATTCGTCGAAGCCTTCCCTATTTGTAACCAAGAGATTAATCTCCCAACAT 1260
 DB 1201 TTGTTCCATTCGTCGAAGCCTTCCCTATTTGTAACCAAGAGATTAATCTCCCAACAT 1260
 QY 1261 CCATCTTACTCANGCAACTTCCATGCAAAACAGCAATATGTTTCTGMAACGATCTATT 1320
 DB 1261 CCATCTTACTCANGCAACTTCCATGCAAAACAGCAATATGTTTCTGMAACGATCTATT 1320
 QY 1321 AAAGATCACAACGCTAGCGTTCTCCCGTACCTTCCCTCTCTGCGGATCTTTT 1380
 DB 1321 AAAGATCACAACGCTAGCGTTCTCCCGTACCTTCCCTCTCTGCGGATCTTTT 1380
 QY 1381 CGTCCACACGCAATG 1394
 DB 1381 CGTCCACACGCAATG 1394

RESULT 2
 AAH76333
 ID AAH76333 standard; DNA; 1394 BP.
 AC AAH76333;
 XX
 XX
 DT 29-OCT-2001 (first entry)
 XX
 XX
 DE Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.
 XX
 KW Ms45; male tissue; regulatory region; transcription; male fertility;
 KW hybrid seed; ds.
 XX
 OS Zea mays.
 XX
 PN W020016097-A2.
 XX
 XX
 PD 23-AUG-2001.
 XX
 PF 13-FEB-2001; 2001WO-US004527.
 XX
 PR 15-FEB-2000; 2000US-00504487.
 XX
 PA (PION-) PIONEER HI-BRED INT INC.
 XX
 PI Albrechten MC, Fox TW, Garnaat CW, Huffman G, Kendal TL;
 XX
 PT WPI; 2001-514772/56.
 DR
 XX
 PT A male tissue-preferred regulatory region comprising nucleotide sequences
 PT essential for initiating transcription of the Ms45 gene useful for
 PT mediating fertility in a male plant.
 PS
 XX
 PS Claim 4; Page 47; 50pp; English.
 XX
 CC The invention provides a male tissue-preferred regulatory region (1)
 CC comprising nucleotide sequences essential for initiating transcription of
 CC the Ms45 gene. A method of mediating male fertility in a plant is
 CC provided that involves introducing an expression vector comprising a
 CC promoter operably linked to (1) into a plant where the exogenous gene
 CC impacts male fertility of the plant and (2) controls expression of the
 CC exogenous gene. A method of producing hybrid seeds is also provided. The

CC Ms45), which encodes a product selected from auxins, rolB and dipteria
 CC toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
 CC and infertile plants
 XX

Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;

Query Match 94.0%; Score 1311; DB 2; Length 1394;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 CCATGGTGTCTCTATGAAAAAGATGATACATGTCTATATCCGTTTCTTAAAGGTC 60
DB 1 CCATGGTGTCTCTATGAAAAAGATGATACATGTCTATATCCGTTTCTTAAAGGTC 60
QY CTCTCTGCTGCTATTACTGATCGGAGTTTACAAAACTTCCAGGGGTCATGAT 120
DB CTCTCTGCTGCTATTACTGATCGGAGTTTACAAAACTTCCAGGGGTCATGAT 120
QY 121 CTCGATGTTCCACTTCTCCACCTCGGTTGCACTTTCTTGATGTCGGTGGTCCAT 180
DB 121 CTCGATGTTCCACTTCTCCACCTCGGTTGCACTTTCTTGATGTCGGTGGTCCAT 180
QY 181 CTGACCGAGGCCCATGACACCTTTGCGGACACCCATCAAGGGCTTTGCGATGGCCCA 240
DB 181 CTGACCGAGGCCCATGACACCTTTGCGGACACCCATCAAGGGCTTTGCGATGGCCCA 240
QY 241 CGAGAGCTATCGGGTGTGATCCAGGGGATATATGTCGCCCAATCGTCACTATA 300
DB 241 CGAGAGCTATCGGGTGTGATCCAGGGGATATATGTCGCCCAATCGTCACTATA 300
QY 301 TTATATCTTTAGATATATTTTAAATTTTGAATAAACAACCTTATACCTTTGCTA 360
DB 301 TTATATCTTTAGATATATTTTAAATTTTGAATAAACAACCTTATACCTTTGCTA 360
QY 361 GGGCTTCAGATGATTTTGGCTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420
DB 361 GGGCTTCAGATGATTTTGGCTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420
QY 421 TCCACTATTGGCACTACCCAGAACAGATTTAAATAAATTAACCAAGTAACTAATC 480
DB 421 TCCACTATTGGCACTACCCAGAACAGATTTAAATAAATTAACCAAGTAACTAATC 480
QY 481 CGAAGCTATCATGTAATGTTTAAAGAAACATCTATTAACAACGATCCTCTTAA 540
DB 481 CGAAGCTATCATGTAATGTTTAAAGAAACATCTATTAACAACGATCCTCTTAA 540
QY 541 CAAGCATTTTGCAGAAAGACAAATATGTTACAGTTTACAAACATCTAAGGCGACA 600
DB 541 CAAGCATTTTGCAGAAAGACAAATATGTTACAGTTTACAAACATCTAAGGCGACA 600
QY 601 TTATATCGAAGGTAAGCTATGAGTTCAGATTTTCTTTTCATCTTGTATTTTGT 660
DB 601 TTATATCGAAGGTAAGCTATGAGTTCAGATTTTCTTTTCATCTTGTATTTTGT 660
QY 661 ATTGTTTTTATATACATTTTCTCTTCAATAGAGTATTTTCTTCGATTTTAT 720
DB 661 ATTGTTTTTATATACATTTTCTCTTCAATAGAGTATTTTCTTCGATTTTAT 720
QY 721 ATGACTATTAAGCTATTTTATATAGACGACGATGTCGTGATTTCTGTTCA 780
DB 721 ATGACTATTAAGCTATTTTATATAGACGACGATGTCGTGATTTCTGTTCA 780
QY 781 TTTCTGATTTTATAGAGTATTTGGAACCTGTTTCTTCAAGAAATTTGATTT 840
DB 781 TTTCTGATTTTATAGAGTATTTGGAACCTGTTTCTTCAAGAAATTTGATTT 840
QY 841 TTTCAAAAAAATTTAGTTTATTTCTTTTATATAAATAGAAAACATTAAGAAAT 900
DB 841 TTTCAAAAAAATTTAGTTTATTTCTTTTATATAAATAGAAAACATTAAGAAAT 900
QY 901 TGGCAGACTAGCCCTGGAATGTTTCCCAATAATTAACAATCACTGTGTATATTT 960
DB 901 TGGCAGACTAGCCCTGGAATGTTTCCCAATAATTAACAATCACTGTGTATATTT 960

```

```

QY 961 GCCAGCCCATATAATTTATTAACCGAAACTGAATCGAGGAAACCAATCGAGCTAT 1020
DB 961 GCCAGCCCATATAATTTATTAACCGAAACTGAATCGAGGAAACCAATCGAGCTAT 1020
QY 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAATCAATTTAAGTATGTCCTCC 1080
DB 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAATCAATTTAAGTATGTCCTCC 1080
QY 1081 TGAGATGTCGGTTTGGCAACATAGCCAGCTAATCATAGCTCATAGTGCCTACGTA 1140
DB 1081 TGAGATGTCGGTTTGGCAACATAGCCAGCTAATCATAGCTCATAGTGCCTACGTA 1140
QY 1141 GGTTCGGAGCTCTGCTGTCATCTCAATGAGCATCTACATGCTTTGTTCAACGTT 1200
DB 1141 GGTTCGGAGCTCTGCTGTCATCTCAATGAGCATCTACATGCTTTGTTCAACGTT 1200
QY 1201 TTGTTTCATCGTCAAGCCTTGCCTATTCTGAACCAAGAGATACCTACTCCAAACAT 1260
DB 1201 TTGTTTCATCGTCAAGCCTTGCCTATTCTGAACCAAGAGATACCTACTCCAAACAT 1260
QY 1261 CCATCTTACTCATGCAACTTCCATGCAACACGCAATATGTTTCTGMAAC 1311
DB 1261 CCATCTTACTCATGCAACTTCCATGCAACACGCAATATGTTTCTGMAAC 1311

```

RESULT 4

AAH76332
 ID AAH76332 standard; DNA; 1394 BP.

AAH76332;

29-OCT-2001 (first entry)

Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.

Ms45; male tissue; regulatory region; transcription; male fertility;

hybrid seed; ds.

Zea mays.

W0200160997-A2.

23-AUG-2001.

13-FEB-2001; 2001WO-US004527.

15-FEB-2000; 2000US-00504487.

(PION-) PIONEER HI-BRED INT INC.

Albertain MC, Fox TW, Garmaat CW, Huffman G, Kendall TL;

WPI; 2001-514772/56.

A male tissue-preferred regulatory region comprising nucleotide sequences

essential for initiating transcription of the MS45 gene useful for

mediating fertility in a male plant.

Claim 4; Page 46; 50pp; English.

The invention provides a male tissue-preferred regulatory region (1)

comprising nucleotide sequences essential for initiating transcription of

the MS45 gene. A method of mediating male fertility in a plant is

provided that involves introducing an expression vector comprising a

promoter operably linked to (1) into a plant where the exogenous gene

impacts male fertility of the plant and (1) controls expression of the

exogenous gene. A method of producing hybrid seeds is also provided. The

present sequence represents a nucleic acid sequence encoding an Ms45 male

tissue preferred regulatory region from Z. mays

Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;

Query Match 94.0%; Score 1311; DB 5; Length 1394;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCATGAGTGTCTCTATGAAAAAGATGATACAAATGTCTATATCCGTTTCTTAAGGCTCC 60
 Db 1 CCATGAGTGTCTCTATGAAAAAGATGATACAAATGTCTATATCCGTTTCTTAAGGCTCC 60

Qy 61 CTTCTTCTGCTTATTAAGCTGAATCGGGGTTTACAAAACTTCCAGGGTGATGAT 120
 Db 61 CTTCTTCTGCTTATTAAGCTGAATCGGGGTTTACAAAACTTCCAGGGTGATGAT 120

Qy 121 CTCATGTTCCACTTCTCCGACCTGGGTTGCACTTCTTGATGTGGGTTCCAT 180
 Db 121 CTCATGTTCCACTTCTCCGACCTGGGTTGCACTTCTTGATGTGGGTTCCAT 180

Qy 181 CTGACCGAGGCCCATGACACACCTTTGGGACCCATCAAGGGCTTTGSGATGGCCA 240
 Db 181 CTGACCGAGGCCCATGACACACCTTTGGGACCCATCAAGGGCTTTGSGATGGCCA 240

Qy 241 CGAGAGCTATCGGCTGGTGATCGAGGGATATATGTCCCAATGTCACCTATA 300
 Db 241 CGAGAGCTATCGGCTGGTGATCGAGGGATATATGTCCCAATGTCACCTATA 300

Qy 301 TTATATTCTTTGATATTAATTTTGGAAAAATPACAACTTATCTTTGTGTA 360
 Db 301 TTATATTCTTTGATATTAATTTTGGAAAAATPACAACTTATCTTTGTGTA 360

Qy 361 GGGGCTCAGCATAGATTTTGGCTTGGGCCCAGAAATGCGAGACAGCAGCTCTAGT 420
 Db 361 GGGGCTCAGCATAGATTTTGGCTTGGGCCCAGAAATGCGAGACAGCAGCTCTAGT 420

Qy 421 TCCACTATTTGGCACTACCCAGAACAGATTTAAAAAATPACAAAGTAACTAATCCACT 480
 Db 421 TCCACTATTTGGCACTACCCAGAACAGATTTAAAAAATPACAAAGTAACTAATCCACT 480

Qy 481 CGAAGCTATCATGTATGTTTAAAGAAACATCATTAACCAAGCTCTTAAATAA 540
 Db 481 CGAAGCTATCATGTATGTTTAAAGAAACATCATTAACCAAGCTCTTAAATAA 540

Qy 541 CAAGCATATTTCCAAAGAGACAAATTTATTTACAGTTTACAAACATCTAAGAGGACAA 600
 Db 541 CAAGCATATTTCCAAAGAGACAAATTTATTTACAGTTTACAAACATCTAAGAGGACAA 600

Qy 601 TTATATCGAAAGGTAGCTATGACGTTCAAGATTTTCTTTTCATCTTGTATTTGTT 660
 Db 601 TTATATCGAAAGGTAGCTATGACGTTCAAGATTTTCTTTTCATCTTGTATTTGTT 660

Qy 661 ATTGTTTTTATATACATTTCTCTCTTACATATGATGATTTTCTTCCGATTTTATAA 720
 Db 661 ATTGTTTTTATATACATTTCTCTCTTACATATGATGATTTTCTTCCGATTTTATAA 720

Qy 721 ATGACTATAAGCTATTTTATATAAGAGACGATGCTAGATTTCTGTTCAAAATC 780
 Db 721 ATGACTATAAGCTATTTTATATAAGAGACGATGCTAGATTTCTGTTCAAAATC 780

Qy 781 TTTCTGATTTTTTAAAGCTAGTTGGCAACCTGTTTCTTTCAAGAAATTTGATTTT 840
 Db 781 TTTCTGATTTTTTAAAGCTAGTTGGCAACCTGTTTCTTTCAAGAAATTTGATTTT 840

Qy 841 TTCAAAAAAATTAAGTTATTTTCTTTTAAATAATGAAAAACCTTGAAGAAAAATGAGT 900
 Db 841 TTCAAAAAAATTAAGTTATTTTCTTTTAAATAATGAAAAACCTTGAAGAAAAATGAGT 900

Qy 901 TGGCAGACTAGCCCTAGAAATGTTTTCCCAATTAATCACTGTGTAATTAATTTG 960
 Db 901 TGGCAGACTAGCCCTAGAAATGTTTTCCCAATTAATCACTGTGTAATTAATTTG 960

Qy 961 GCCAGCCCCATTAATTTTAAACCGAAACTGAAATGAGCGAAACCAAACTGAGCTAT 1020
 Db 961 GCCAGCCCCATTAATTTTAAACCGAAACTGAAATGAGCGAAACCAAACTGAGCTAT 1020

Qy 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAAATCATGTTTAAATCATGTGCC 1080

Db 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAAATCATGTTTAAATCATGTGCC 1080

Qy 1081 TGAGATGTGCGGTTTGGCAACGATAGCACCGTAATCATATGATAGTGCTTACGTCA 1140
 Db 1081 TGAGATGTGCGGTTTGGCAACGATAGCACCGTAATCATATGATAGTGCTTACGTCA 1140

Qy 1141 GGTTCGGACGCTCTCGTGTATCTCACATGGCATCTACTACATGCTTTGTTCAACGTTGCTC 1200
 Db 1141 GGTTCGGACGCTCTCGTGTATCTCACATGGCATCTACTACATGCTTTGTTCAACGTTGCTC 1200

Qy 1201 TTGTTTCATGCTGCAAGCCCTGCTATTTGAAACCAAGAGATACCTACCCGAAACAT 1260
 Db 1201 TTGTTTCATGCTGCAAGCCCTGCTATTTGAAACCAAGAGATACCTACCCGAAACAT 1260

Qy 1261 CCATCTTACTCATGCAACTTCCATGCAACACCGACATATGTTCTCGAAC 1311
 Db 1261 CCATCTTACTCATGCAACTTCCATGCAACACCGACATATGTTCTCGAAC 1311

RESULT 5
 AAH76340
 ID AAH76340 standard; DNA; 255 BP.
 XX
 AC AAH76340;
 XX
 DT 29-OCT-2001 (first entry)
 XX
 DE Z. may5 Ms45 promoter fragment.
 XX
 KW Ms45; male tissue; regulatory region; transcription; male fertility;
 KW hybrid seed; promoter; de.
 XX
 OS Zea mays.
 XX
 PN NC0200160997-A2.
 PD 23-AUG-2001.
 XX
 PF 13-FEB-2001; 2001MO-US004527.
 PR 15-FEB-2000; 2000US-00504487.
 PA (PION-) PIONEER HI-BRED INT INC.
 XX
 PI Albrechtsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
 XX
 DR WPI; 2001-514772/56.
 XX
 PT A male tissue-preferred regulatory region comprising nucleotide sequences
 PT essential for initiating transcription of the Ms45 gene useful for
 PT mediating fertility in a male plant.
 XX
 PS Example 5; Fig 8; 50pp; English.
 XX
 CC The invention provides a male tissue-preferred regulatory region (I)
 CC comprising nucleotide sequences essential for initiating transcription of
 CC the Ms45 gene. A method of mediating male fertility in a plant is
 CC provided that involves introducing an expression vector comprising a
 CC promoter operably linked to (I) into a plant where the exogenous gene
 CC impacts male fertility of the plant and (I) controls expression of the
 CC exogenous gene. A method of producing hybrid seeds is also provided. The
 CC present sequence represents a Z. mays Ms45 promoter fragment
 XX
 SQ Sequence 255 BP; 59 A; 86 C; 39 G; 71 T; 0 U; 0 Other;

Query Match 17.2%; Score 240; DB 5; Length 255;
 Best Local Similarity 100.0%; Pred. No. 9.4e-105; Indels 0; Gaps 0;
 Matches 240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1155 CGTGTATCTCATGAGCATGATGATGCTTGTCAACGTTGCTGTTGCATGCTCC 1214
 Db 15 CGTGTATCTCATGAGCATGATGATGCTTGTCAACGTTGCTGTTGCATGCTCC 74

QY 1215 AACCTTGCTATTTGAAACCAAGAGATACCTACTCCAAACATTCATCTCATG 1274
 |||||
 DB 75 AACCTTGCTATTTGAAACCAAGAGATACCTACTCCAAACATTCATCTCATG 134
 |||||
 QY 1275 CAATCTTCATGCAAAACGACACATATGTTCTTGAAACAGATCTATTAAAGATCAACAG 1334
 |||||
 DB 135 CAATCTTCATGCAAAACGACACATATGTTCTTGAAACAGATCTATTAAAGATCAACAG 194
 |||||
 QY 1335 CTAGCGTTCTCCGCTAGCTTCCCTCTCTCTGCGAGATCTTTTCGTCACACAGG 1394
 |||||
 DB 195 CTAGCGTTCTCCGCTAGCTTCCCTCTCTCTGCGAGATCTTTTCGTCACACAGG 254
 |||||
 RESULT 6
 AAH76334
 ID AAH76334 standard; DNA; 158 BP.
 AC AAH76334;
 XX
 DT 29-OCT-2001 (first entry)
 XX
 DE 2. mayb Ms45 male tissue-preferred regulatory region fragment.
 XX
 KW Ms45; male tissue; regulatory region; transcription; male fertility;
 XX hybrid seed; ds.
 XX
 OS Zea mays.
 XX
 PN WO200160997-A2.
 XX
 PD 23-AUG-2001.
 XX
 PF 13-FEB-2001; 2001WO-US004527.
 XX
 PR 15-FEB-2000; 2000US-00504487.
 XX
 PA (PION-) PIONEER HI-BRED INT INC.
 XX
 PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
 XX
 DR WPI; 2001-514772/56.
 XX
 XX A male tissue-preferred regulatory region comprising nucleotide sequences
 PT essential for initiating transcription of the Ms45 gene useful for
 PT mediating fertility in a male plant.
 XX
 PS Claim 5; Page 47; 50pp; English.
 XX
 CC The invention provides a male tissue-preferred regulatory region (I)
 CC comprising nucleotide sequences essential for initiating transcription of
 CC the Ms45 gene. A method of mediating male fertility in a plant is
 CC provided that involves introducing an expression vector comprising a
 CC promoter operably linked to (i) into a plant where the exogenous gene
 CC impacts male fertility of the plant and (i) controls expression of the
 CC exogenous gene. A method of producing hybrid seeds is also provided. The
 CC present sequence represents a DNA fragment -38 to -195 bases upstream of
 CC the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
 CC nucleotide sequence
 CC
 SQ Sequence 158 BP; 41 A; 50 C; 21 G; 46 T; 0 U; 0 Other;
 XX
 XX
 Query Match 8.0%; Score 111; DB 5; Length 158;
 Best Local Similarity 100.0%; Pred. No. 1.1e-42;
 Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1201 TTGTTCCATGCTCAAGCTTGGCTATTGTAACCAAGAGATACCTACTCCAAACAT 1260
 |||||
 DB 48 TTGTTCCATGCTCAAGCTTGGCTATTGTAACCAAGAGATACCTACTCCAAACAT 107
 |||||
 QY 1261 CCATCTTACATGCAAACTTCATGCAACGCAACATATGTTCTCTGAAC 1311
 |||||
 DB 108 CCATCTTACATGCAAACTTCATGCAACGCAACATATGTTCTCTGAAC 158
 |||||

RESULT 7
 AAH76336
 ID AAH76336 standard; DNA; 50 BP.
 AC AAH76336;
 XX
 DT 29-OCT-2001 (first entry)
 XX
 DE 2. mayb Ms45 male tissue-preferred regulatory region fragment.
 XX
 KW Ms45; male tissue; regulatory region; transcription; male fertility;
 XX hybrid seed; ds.
 XX
 OS Zea mays.
 XX
 PN WO200160997-A2.
 XX
 PD 23-AUG-2001.
 XX
 PF 13-FEB-2001; 2001WO-US004527.
 XX
 PR 15-FEB-2000; 2000US-00504487.
 XX
 PA (PION-) PIONEER HI-BRED INT INC.
 XX
 PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
 XX
 DR WPI; 2001-514772/56.
 XX
 XX A male tissue-preferred regulatory region comprising nucleotide sequences
 PT essential for initiating transcription of the Ms45 gene useful for
 PT mediating fertility in a male plant.
 XX
 PS Claim 5; Page 47; 50pp; English.
 XX
 CC The invention provides a male tissue-preferred regulatory region (I)
 CC comprising nucleotide sequences essential for initiating transcription of
 CC the Ms45 gene. A method of mediating male fertility in a plant is
 CC provided that involves introducing an expression vector comprising a
 CC promoter operably linked to (i) into a plant where the exogenous gene
 CC impacts male fertility of the plant and (i) controls expression of the
 CC exogenous gene. A method of producing hybrid seeds is also provided. The
 CC present sequence represents a DNA fragment -72 to -111 bases upstream of
 CC the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
 CC nucleotide sequence
 CC
 SQ Sequence 50 BP; 18 A; 17 C; 5 G; 10 T; 0 U; 0 Other;
 XX
 XX
 Query Match 3.6%; Score 50; DB 5; Length 50;
 Best Local Similarity 100.0%; Pred. No. 2.4e-13;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1229 CTGAACCAAGAGATACCTACTCCCAACATTCATCTTACATGCAAC 1278
 |||||
 DB 1 CTGAACCAAGAGATACCTACTCCCAACATTCATCTTACATGCAAC 50
 |||||
 RESULT 8
 AAH76337
 ID AAH76337 standard; DNA; 40 BP.
 AC AAH76337;
 XX
 DT 29-OCT-2001 (first entry)
 XX
 DE 2. mayb Ms45 male tissue-preferred regulatory region fragment.
 XX
 KW Ms45; male tissue; regulatory region; transcription; male fertility;
 XX hybrid seed; ds.
 XX
 OS Zea mays.
 XX

XX WO200160997-A2.
PN 23-AUG-2001.
XX 13-FEB-2001; 2001WO-US004527.
PF 15-FEB-2000; 2000US-00504487.
XX (PION-) PIONEER HI-BRED INT INC.
PA
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL,
DR WPI; 2001-514772/56.
XX
XX A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the MS45 gene useful for
PS mediating fertility in a male plant.
PS Claim 14; Page 32; 50pp; English.
XX
XX The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the MS45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a DNA fragment upstream of the TATA box of a
CC Z. mays Ms45 male-tissue preferred regulatory region nucleotide sequence
XX
SQ Sequence 40 BP; 14 A; 14 C; 3 G; 9 T; 0 U; 0 Other;
XX
XX Query Match 2.9%; Score 40; DB 5; Length 40;
Best Local Similarity 100.0%; Pred. No. 1.6e-08;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
Qy 1239 AGGATACCTACTCCCAACATCATCTTACTCATGCAAC 1278
Db 1 AGGATACCTACTCTCCCAACATCATCTTACTCATGCAAC 40
XX
RESULT 9
AAH76335
ID AAH76335 standard; DNA; 30 BP.
XX
XX AAH76335;
AC
XX
XX 29-OCT-2001 (first entry)
DT
XX
XX Z. mays Ms45 male tissue-preferred regulatory region fragment.
DE
XX
XX Ms45; male tissue; regulatory region; transcription; male fertility;
KM hybrid seed; ds.
XX
XX Zea mays.
OS
XX
XX WO200160997-A2.
PN
XX
XX 23-AUG-2001.
PD
XX
XX 13-FEB-2001; 2001WO-US004527.
PF
XX
XX 15-FEB-2000; 2000US-00504487.
PR
XX
XX (PION-) PIONEER HI-BRED INT INC.
PA
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL,
DR WPI; 2001-514772/56.
XX
XX A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the MS45 gene useful for

PT mediating fertility in a male plant.
XX
XX Claim 5; Page 47; 50pp; English.
PS
XX
XX The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the MS45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a DNA fragment -152 to -181 bases upstream of
CC the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
XX
SQ Sequence 30 BP; 4 A; 9 C; 5 G; 12 T; 0 U; 0 Other;
XX
XX Query Match 2.2%; Score 30; DB 5; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
Qy 1179 CATGCTGTTCACCGTTCGTTGTTCCA 1208
Db 1 CATGCTGTTCACCGTTCGTTGTTCCA 30
XX
RESULT 10
ABL40437/c
ID ABL40437 standard; DNA; 849 BP.
XX
XX ABL40437;
AC
XX
XX 10-JUN-2002 (first entry)
DT
XX
XX Sunflower chitinase promoter sequence.
DE
XX
XX Sunflower; chitinase; lipid transfer protein; LTP; transcription;
KM sclerotinia; genetic engineering; disease resistance; insecticide;
KM antifungal; viricide; nematocide; antimicrobial; antibacterial;
KM gene therapy; gene; promoter; ds.
XX
XX Helianthus annuus.
OS
XX
XX
XX Key Location/Qualifiers
FH CHAT_signal 723..726
FT /*tag= a
FT TATA_signal 807..811
FT /*tag= b
XX
XX WO200214502-A2.
PN
XX
XX 21-FEB-2002.
PD
XX
XX 08-AUG-2001; 2001WO-US041629.
PF
XX
XX 11-AUG-2000; 2000US-0224603P.
PR
XX
XX 07-AUG-2001; 2001US-00923844.
PR
XX
XX (PION-) PIONEER HI-BRED INT INC.
PA
XX
XX Bao Z, Lu G;
PI
XX
XX WPI; 2002-269194/31.
DR
XX
XX New genes and promoters from Helianthus annuus, for genetically
PT manipulating plants to enhance disease resistance or resistance to
PT pathogen, and for producing proteins for controlling plant diseases
PT caused by these pathogens.
XX
XX Claim 1; Fig 4; 89pp; English.
PS
XX
XX The invention relates to sclerotinia-inducible genes and promoters capable
CC of initiating transcription from sunflower. The polynucleotides are

CC useful for genetically manipulating plants to enhance disease resistance,
CC or resistance to insect, fungal pathogens, viruses or nematodes. The
CC nucleotide sequences are also useful as genetic markers in disease-
CC resistance breeding programs. The polynucleotides are also useful for
CC producing proteins useful in compositions. The compositions are useful
CC reducing or eliminating damage to plants caused by plant pathogens. The
CC compositions find use in agricultural and pharmaceutical compositions as
CC antifungal and antimicrobial agent. For agricultural purposes, the
CC compositions may be used in sprays for control of plant disease. As
CC pharmaceutical compositions, the agents are useful as antibacterial and
CC antimicrobial treatments. The present sequence represents the promoter
CC fragment of the sunflower chitinase gene
XX

SEQ Sequence 849 BP; 295 A; 139 C; 135 G; 280 T; 0 U; 0 Other;

Query Match 1.6%; Score 23; DB 6; Length 849;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 854 AGTTATTTCTCTTATAAAT 876
|||||
DB 556 AGTTATTTCTCTTATAAAT 534

RESULT 11
ABV56505
ID ABV56505 standard; cDNA; 333 BP.
XX

ABV56505;

17-SEP-2002 (first entry)

Human prostate expression marker cDNA 56496.

Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KM pharmacogenomic marker; gene; ss.

Homo sapiens.

MO200160860-A2.

23-AUG-2001.

20-FEB-2001; 2001WO-US005171.

17-FEB-2000; 2000US-0183119P.

16-MAR-2000; 2000US-0189862P.

25-MAY-2000; 2000US-0207454P.

09-JUN-2000; 2000US-0211314P.

18-JUL-2000; 2000US-0219007P.

13-DEC-2000; 2000US-0255281P.

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

Schlegel R, Endege WO, Monahan DE;

WPI; 2001-662795/76.

Novel isolated nucleic acid molecule associated with cancerous state of

prostate cells and correlating with presence of prostate cancer, useful

for detecting presence of prostate cancer, stage of prostate cancer.

Claim 1; Page 10900; 11750pp; English.

The invention relates to an isolated nucleic acid molecule (1) comprising
a nucleotide sequence given in Tables 1-9 (ABV0010-ABV6213) of the
specification or its complement. (1) is useful for: (a) assessing whether
a patient is afflicted with prostate cancer; (b) monitoring the
progression of prostate cancer in a patient; (c) assessing the efficacy
of a test compound to inhibit prostate cancer in a patient; (d) assessing
the efficacy of a therapy for inhibiting prostate cancer in a patient;
(e) selecting a composition for inhibiting prostate cancer in a patient;
(f) assessing the prostate cell carcinogenic potential of a compound; (g)

CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC; (1) is also useful as a pharmacodynamic or pharmacogenomic marker
XX

SEQ Sequence 333 BP; 75 A; 69 C; 46 G; 143 T; 0 U; 0 Other;

Query Match 1.5%; Score 21; DB 5; Length 333;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 93 TTACAAAAAAGCTTCACGGGT 113
|||||
DB 89 TTACAAAAAAGCTTCACGGGT 109

RESULT 12
ABA58867/C
ID ABA58867 standard; DNA; 459 BP.
XX

ABA58867;

01-FEB-2002 (first entry)

Human foetal liver single exon nucleic acid probe #7172.

Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

Homo sapiens.

WO200157277-A2.

09-AUG-2001.

30-JAN-2001; 2001WO-US000669.

04-FEB-2000; 2000US-0180312P.

26-MAY-2000; 2000US-0207456P.

30-JUN-2000; 2000US-00608408.

03-AUG-2000; 2000US-00632366.

21-SEP-2000; 2000US-0234687P.

27-SEP-2000; 2000US-0236359P.

04-OCT-2000; 2000GB-00024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;

WPI; 2001-483447/52.

Human genome-derived single exon nucleic acid probes useful for analyzing

gene expression in human fetal liver.

Claim 1; SEQ ID NO 7172; 639pp + Sequence Listing; English.

The invention relates to a single exon nucleic acid probe for measuring

human gene expression in a sample derived from human foetal liver. The

single exon nucleic acid probes may be used for predicting, measuring and

displaying gene expression in samples derived from human fetal liver. The

present sequence is a single exon nucleic acid probe of the invention.

Note: The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pct_sequences

Sequence 459 BP; 86 A; 149 C; 70 G; 154 T; 0 U; 0 Other;

Query Match 1.5%; Score 21; DB 4; Length 459;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1037 AAAGGAGAGAGAGAGAGA 1057
|||||
DB 118 AAAGGAGAGAGAGAGAGA 98

```
RESULT 13
AAI38581/c
ID AAI38581 standard; DNA; 459 BP.
XX
AC AAI38581;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #7267 used to measure gene expression in human placenta sample.
XX
KM Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000663.
XX
PR 04-FEB-2000; 2000US-0180312P.
XX
PR 26-MAY-2000; 2000US-0207456P.
XX
PR 30-JUN-2000; 2000US-00608408.
XX
PR 03-AUG-2000; 2000US-00632366.
XX
PR 21-SEP-2000; 2000US-0234687P.
XX
PR 27-SEP-2000; 2000US-0236359P.
XX
PR 04-OCT-2000; 2000GB-00024263.
XX
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DX WPI; 2001-488897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human placenta.
XX
PS Claim 25; SEQ ID NO 7267; 654bp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP).
XX
CC The present sequence is one such probe. The probes are useful for
XX producing a microarray for predicting, measuring and displaying gene
XX expression in samples derived from human placenta. The probes are useful
XX for antenatal diagnosis of human genetic disorders
XX
SQ Sequence 459 BP; 86 A; 149 C; 70 G; 154 T; 0 U; 0 Other;

Query Match          1.5%; Score 21; DB 4; Length 459;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1037 AAAGGAGAGAGAGAGAGA 1057
      |||||
Db      118 AAAGGAGAGAGAGAGAGA 98

RESULT 14
AAK32768/c
ID AAK32768 standard; DNA; 459 BP.
XX
AC AAK32768;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed single exon probe SEQ ID NO: 7325.
XX
KM Human; bone marrow expressed exon; gene expression analysis; probe;
XX microarray; cancer; leukemia; lymphoma; myeloma; ss.
XX
OS Homo sapiens.
XX
```

```
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000668.
XX
PR 04-FEB-2000; 2000US-0180312P.
XX
PR 26-MAY-2000; 2000US-0207456P.
XX
PR 30-JUN-2000; 2000US-00608408.
XX
PR 03-AUG-2000; 2000US-00632366.
XX
PR 21-SEP-2000; 2000US-0234687P.
XX
PR 27-SEP-2000; 2000US-0236359P.
XX
PR 04-OCT-2000; 2000GB-00024263.
XX
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DX WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human bone marrow.
XX
PS Example 4; SEQ ID NO 7325; 658bp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX bone marrow. They can be used to measure gene expression in bone marrow
XX samples, which may enable the improved diagnosis and treatment of cancers
XX such as lymphoma, leukemia and myeloma. The present sequence is one of
XX the probes of the invention
XX
SQ Sequence 459 BP; 86 A; 149 C; 70 G; 154 T; 0 U; 0 Other;

Query Match          1.5%; Score 21; DB 4; Length 459;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1037 AAAGGAGAGAGAGAGA 1057
      |||||
Db      118 AAAGGAGAGAGAGAGA 98

RESULT 15
AAK07027/c
ID AAK07027 standard; DNA; 459 BP.
XX
AC AAK07027;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe SEQ ID NO: 7018.
XX
KM Human; brain expressed exon; gene expression analysis; probe; microarray;
XX Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
XX ss.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000667.
XX
PR 04-FEB-2000; 2000US-0180312P.
XX
PR 26-MAY-2000; 2000US-0207456P.
XX
PR 30-JUN-2000; 2000US-00608408.
XX
PR 03-AUG-2000; 2000US-00632366.
XX
PR 21-SEP-2000; 2000US-0234687P.
XX
PR 27-SEP-2000; 2000US-0236359P.
XX
PR 04-OCT-2000; 2000GB-00024263.
XX
```

PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-483446/52.
 XX
 PT Single exon nucleic acid probes for analyzing gene expression in human
 PT brains.
 XX
 PS Example 4; SEQ ID NO 7018; 650pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is one of the probes of the
 CC invention
 XX
 SQ Sequence 459 BP; 86 A; 149 C; 70 G; 154 T; 0 U; 0 Other;

Query Match 1.5%; Score 21; DB 4; Length 459;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1037 AAAGGAGAGAGAGAGAGA 1057
 |||||
 Db 118 AAAGGAGAGAGAGAGAGA 98

Search completed: September 16, 2005, 13:08:40
 Job time : 579.5 secs

QY 1 CCATGCTCTCTATGAAAAAGATGATCATATGTCATATCCGTTTCTTAAAGGCTC 60
 DB 1 CCATGCTCTCTATGAAAAAGATGATCATATGTCATATCCGTTTCTTAAAGGCTC 60
 QY 61 CTTCCTTCTGCTTATTAATGATCGAATCGGAGTTTACAAAACCTTCCAGGCTGATAT 120
 DB 61 CTTCCTTCTGCTTATTAATGATCGAATCGGAGTTTACAAAACCTTCCAGGCTGATAT 120
 QY 121 CTTCATGTTCCATCTTCCAGCTCGGCTGCAATTTCTTGAATGTCGGTGGTCCAT 180
 DB 121 CTTCATGTTCCATCTTCCAGCTCGGCTGCAATTTCTTGAATGTCGGTGGTCCAT 180
 QY 181 CTGACCGAGGCCCATGACACCTTTGCGGACACCATCAAGGSCCTTTCGATGCGCCA 240
 DB 181 CTGACCGAGGCCCATGACACCTTTGCGGACACCATCAAGGSCCTTTCGATGCGCCA 240
 QY 241 CGAGAGCTATCGGGTGTGTGATCCAGGGGATATATGTCCTCCCAATGTCACCTATA 300
 DB 241 CGAGAGCTATCGGGTGTGTGATCCAGGGGATATATGTCCTCCCAATGTCACCTATA 300
 QY 301 TTATTTATCTTTAGATATATTTTGAAGAAAAAACAACCTTATCTTTGTGTA 360
 DB 301 TTATTTATCTTTAGATATATTTTGAAGAAAAAACAACCTTATCTTTGTGTA 360
 QY 361 GGGCCTCAGCATGATTTTCGTTAAGGCCCAGAAATGCGAGGACGACCATGCTATG 420
 DB 361 GGGCCTCAGCATGATTTTCGTTAAGGCCCAGAAATGCGAGGACGACCATGCTATG 420
 QY 421 TTCACATTTGGCATCTACCCGAAACAGATTTAAAAAATAACCAAGTAACTATGCACT 480
 DB 421 TTCACATTTGGCATCTACCCGAAACAGATTTAAAAAATAACCAAGTAACTATGCACT 480
 QY 481 CGAAGCTATCATGTAATGTTTAAAGAAACATCATTTAAACACGATCCCTTAAAAAA 540
 DB 481 CGAAGCTATCATGTAATGTTTAAAGAAACATCATTTAAACACGATCCCTTAAAAAA 540
 QY 541 CAAGCATATTTGAAAGAGACAAATTAATGTTACAGTTTACAAACATCTAAGACGACAA 600
 DB 541 CAAGCATATTTGAAAGAGACAAATTAATGTTACAGTTTACAAACATCTAAGACGACAA 600
 QY 601 TTATATCGAAAGGTAAAGCTATGACGTTCAATTTTCTTTTCAATCTTGTATATTTGTT 660
 DB 601 TTATATCGAAAGGTAAAGCTATGACGTTCAATTTTCTTTTCAATCTTGTATATTTGTT 660
 QY 661 ATGTTTTTATATACATTTTCTTCTCTTACATAGAGTATTTCTTCCGATTTTAA 720
 DB 661 ATGTTTTTATATACATTTTCTTCTCTTACATAGAGTATTTCTTCCGATTTTAA 720
 QY 721 ATGACTATAAAGTCAATTTTATAAAGACGACATGCTAGATTTCTGTTCAAAAAATC 780
 DB 721 ATGACTATAAAGTCAATTTTATAAAGACGACATGCTAGATTTCTGTTCAAAAAATC 780
 QY 781 TTTCTGATTTTAAAGAGTATTTGGCAACCTGTTTCTTCAAGAAATTTGATTTT 840
 DB 781 TTTCTGATTTTAAAGAGTATTTGGCAACCTGTTTCTTCAAGAAATTTGATTTT 840
 QY 841 TTCAAAAAAATAGTTTATTTCTCTTATATAAATAGAAAAACCTTGAAGATAGAGT 900
 DB 841 TTCAAAAAAATAGTTTATTTCTCTTATATAAATAGAAAAACCTTGAAGATAGAGT 900
 QY 901 TGGCAGACTAGCCCTAGAAATGTTTCCCAATAAATTAACATCACTGTGATATATTTG 960
 DB 901 TGGCAGACTAGCCCTAGAAATGTTTCCCAATAAATTAACATCACTGTGATATATTTG 960
 QY 961 GCGAGGCCCATTAATTTATTTAAACCGAAACTGAATCGAGCGAAACCAATCTGACCTAT 1020
 DB 961 GCGAGGCCCATTAATTTATTTAAACCGAAACTGAATCGAGCGAAACCAATCTGACCTAT 1020
 QY 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAATCACTTTTAAAGCTATGCTCC 1080
 DB 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAATCACTTTTAAAGCTATGCTCC 1080
 QY 1081 TGAGATGTGGCTTTGGCAACATAGCACCGTAAATCATAGCTCATAGGTGCTTACGTCA 1140

DB 1081 TGAGATGTGGCTTTGGCAACATAGCACCGTAAATCATAGCTCATAGGTGCTTACGTCA 1140
 QY 1141 GGTGGGACGCTCTGCTGATCTCATGACATGACATACCTGTTTCAACCGTTGCTC 1200
 DB 1141 GGTGGGACGCTCTGCTGATCTCATGACATGACATACCTGTTTCAACCGTTGCTC 1200
 QY 1201 TTGTTCCATGTCGCAAGCCTTTCCTATTCTGAACCAAGAGATACCTATCTCCCAACAT 1260
 DB 1201 TTGTTCCATGTCGCAAGCCTTTCCTATTCTGAACCAAGAGATACCTATCTCCCAACAT 1260
 QY 1261 CCATCTTACTCATGCAACTTCCATGCAACGACATATGTTTCTGAAACAGATCTATT 1320
 DB 1261 CCATCTTACTCATGCAACTTCCATGCAACGACATATGTTTCTGAAACAGATCTATT 1320
 QY 1321 AAAAGTCAACAAAGCTAGGTTCTCCGCTACTTCCCTCTCTCTGCGCATCTTTT 1380
 DB 1321 AAAAGTCAACAAAGCTAGGTTCTCCGCTACTTCCCTCTCTCTGCGCATCTTTT 1380
 QY 1381 CGTCCACGACCATG 1394
 DB 1381 CGTCCACGACCATG 1394

RESULT 2
 US-08-880-499-1
 ; Sequence 1, Application US/08880499
 ; Patent No. 6037523
 ; GENERAL INFORMATION:
 ; APPLICANT: Albertson, Marc C.
 ; APPLICANT: Carl, Tim W.
 ; APPLICANT: Fox, Garnet W.
 ; APPLICANT: Hufman, Gary A.
 ; APPLICANT: Kendall, Timmy L.
 ; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION
 ; TITLE OF INVENTION: AND METHOD OF USING SAME
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
 ; STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.
 ; CITY: Johnston
 ; STATE: Iowa
 ; COUNTRY: USA
 ; ZIP: 50131
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/880 499
 ; FILING DATE: CONCURRENTLY HERewith
 ; CLASSIFICATION: 800
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Sweeney, Patricia A.
 ; REGISTRATION NUMBER: 32,733
 ; REFERENCE/DOCKET NUMBER: 0578
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (515) 248-4800
 ; TELEFAX: (515) 248-4844
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1394 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; US-08-880-499-1
 Query Match 94.0%; Score 1311; DB 3; Length 1394;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 CCAATGCTCTATGAAAAAGATGATACAAATGCTCTATATCCGTTTTCTTAGGGTCC 60
Db 1 CCAATGCTCTATGAAAAAGATGATACAAATGCTCTATATCCGTTTTCTTAGGGTCC 60
Qy 61 CTTCTTCGCTTATATGACTGATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120
Db 61 CTTCTTCGCTTATATGACTGATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120
Qy 121 CTGATGTTGCACTTCTCCCACTCGGCTTGACATTTCTTGATGTCGGTGGTCCAT 180
Db 121 CTGATGTTGCACTTCTCCCACTCGGCTTGACATTTCTTGATGTCGGTGGTCCAT 180
Qy 181 CTGACGAGGCGCATGACACCTTTGGGACACCCCAAGGGCCCTTGGATGGGCCA 240
Db 181 CTGACGAGGCGCATGACACCTTTGGGACACCCCAAGGGCCCTTGGATGGGCCA 240
Qy 241 CGAGACGTATCGGTCGTGGTGAATCCAGGGATATATGTCGCCCAACATGTCACCTATA 300
Db 241 CGAGACGTATCGGTCGTGGTGAATCCAGGGATATATGTCGCCCAACATGTCACCTATA 300
Qy 301 TTATATTTCTTTAGATATTAATTTATTTTGGAAAAATTAACAACTTATACCTTTGTGA 360
Db 301 TTATATTTCTTTAGATATTAATTTATTTTGGAAAAATTAACAACTTATACCTTTGTGA 360
Qy 361 GGGGCTCAGATGATTTTGGCTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420
Db 361 GGGGCTCAGATGATTTTGGCTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420
Qy 421 TCCACTATTTGGCACTACCCAGAACAGATTTAAAAAAATTAACAAAGTAATCACTACT 480
Db 421 TCCACTATTTGGCACTACCCAGAACAGATTTAAAAAAATTAACAAAGTAATCACTACT 480
Qy 481 CGAAGCTATCATGTAATGTTTAAAGAAACATCTATTAACCAAGATCTCTTAAAAAA 540
Db 481 CGAAGCTATCATGTAATGTTTAAAGAAACATCTATTAACCAAGATCTCTTAAAAAA 540
Qy 541 CAAGCATATTTGCAAGAGAGCAAAATTTGTTACAGTTTAAACATCTAGACGACAA 600
Db 541 CAAGCATATTTGCAAGAGAGCAAAATTTGTTACAGTTTAAACATCTAGACGACAA 600
Qy 601 TTATATCGAAGAGTGAAGCTATGACGTTCAATTTTCTTTTCAATTTCTTTATTTGTT 660
Db 601 TTATATCGAAGAGTGAAGCTATGACGTTCAATTTTCTTTTCAATTTCTTTATTTGTT 660
Qy 661 ATTGTTTTTATATACATTTTCTTCTTAAATAGATGATTTTCTTCCGATTTTATAAA 720
Db 661 ATTGTTTTTATATACATTTTCTTCTTAAATAGATGATTTTCTTCCGATTTTATAAA 720
Qy 721 ATGACTATTAAGCTATTTTATATTAAGAGACGATGTCGTAGATTTCTCTTAAAAATC 780
Db 721 ATGACTATTAAGCTATTTTATATTAAGAGACGATGTCGTAGATTTCTCTTAAAAATC 780
Qy 781 TTTCTGATTTTTTAAAGAGCTAGTTTGGCAACCTGTTTCTTCAAAAGATTTGATTTT 840
Db 781 TTTCTGATTTTTTAAAGAGCTAGTTTGGCAACCTGTTTCTTCAAAAGATTTGATTTT 840
Qy 841 TTCAAAAAAATAGTTTATTTTCTTTTAAATTAAGAAACACTTAGAAAAATAGAGT 900
Db 841 TTCAAAAAAATAGTTTATTTTCTTTTAAATTAAGAAACACTTAGAAAAATAGAGT 900
Qy 901 TGGCAGCTAGCCCTAGAAATGTTTCCCAATTAATTCATCTGCTGATATTAATTTG 960
Db 901 TGGCAGCTAGCCCTAGAAATGTTTCCCAATTAATTCATCTGCTGATATTAATTTG 960
Qy 961 GCCAGCCCATTAATTTTAAACCGAATCTGAATCGAGCAAAACCAATCTGAGTAT 1020
Db 961 GCCAGCCCATTAATTTTAAACCGAATCTGAATCGAGCAAAACCAATCTGAGTAT 1020
Qy 1021 TTCTCTAGATTAAGTAAAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Db 1021 TTCTCTAGATTAAGTAAAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080

```

```

Qy 1081 TGAGATGTGGGTTTGGCAACGATAGCAACCGTAATCATAGCTCATAGGTCCTAGCTCA 1140
Db 1081 TGAGATGTGGGTTTGGCAACGATAGCAACCGTAATCATAGCTCATAGGTCCTAGCTCA 1140
Qy 1141 GGTTCGGACGCTCTCGTGCATCTCATGTCGATGATGATGATGATGATGATGATGATGAT 1200
Db 1141 GGTTCGGACGCTCTCGTGCATCTCATGTCGATGATGATGATGATGATGATGATGATGAT 1200
Qy 1201 TTGTTCCATGCTGCAAGCGCTTGTCTTCTGAAACCAAGAGATCTACTTCCCAACAT 1260
Db 1201 TTGTTCCATGCTGCAAGCGCTTGTCTTCTGAAACCAAGAGATCTACTTCCCAACAT 1260
Qy 1261 CCATCTTACTCATGCAACTTCCATGCAACGACGACATATGTTCTCTGAAC 1311
Db 1261 CCATCTTACTCATGCAACTTCCATGCAACGACGACATATGTTCTCTGAAC 1311

RESULT 3
US-09-949-016-13563
; Sequence 13563: Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13563
; LENGTH: 50263
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13563

Query Match
Best local Similarity 1.5%; Score 21; DB 4; Length 50263;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 861 TTTCTCTTATTAATAATAGAAA 881
Db 40441 TTTCTCTTATTAATAATAGAAA 40461

RESULT 4
US-09-949-016-124481/c
; Sequence 124481: Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 124481
; LENGTH: 601
; TYPE: DNA

```

ORGANISM: Human
US-09-949-016-124481

Query Match 1.4%; Score 20; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 630 GATTTTCTTTTCATCTT 649
|||||
DB 363 GATTTTCTTTTCATCTT 344

RESULT 5
US-09-949-016-147075/c
Sequence 147075, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 147075
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-147075

Query Match 1.4%; Score 20; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 817 TTTCCTTCAAGATTTCGA 836
|||||
DB 353 TTTCCTTCAAGATTTCGA 334

RESULT 6
US-08-410-784A-3/c
Sequence 3, Application US/08410784A
Patent No. 5912413
GENERAL INFORMATION:
APPLICANT: MYERS, ALAN M.
APPLICANT: JAMES, MARTHA G.
TITLE OF INVENTION: ISOLATION OF SUL, A STARCH DEBRANCHING
ENZYME, THE PRODUCT OF THE MAIZE GENE
TITLE OF INVENTION: SUGARY 1
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Weingarten, Schurgin, Gagnebin and Hayes LLP
STREET: Ten Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/410,784A
FILING DATE: 24-MAR-1995
CLASSIFICATION: 800

PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Heine, Ph.D., Holiday C
REGISTRATION NUMBER: 34,346
REFERENCE/DOCKET NUMBER: ISU-002XX
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-2290
TELEFAX: 617-451-0313
TELEX:

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2523 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
US-08-410-784A-3

Query Match 1.4%; Score 20; DB 2; Length 2523;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 849 AATTAGTTTATTTCTCTT 868
|||||
DB 206 AATTAGTTTATTTCTCTT 187

RESULT 7
US-09-949-016-15250
Sequence 15250, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15250
LENGTH: 21679
TYPE: DNA
ORGANISM: Human
US-09-949-016-15250

Query Match 1.4%; Score 20; DB 4; Length 21679;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 630 GATTTTCTTTTCATCTT 649
|||||
DB 4880 GATTTTCTTTTCATCTT 4899

RESULT 8
US-09-949-016-12426/c
Sequence 12426, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.


```

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12426
; LENGTH: 65424
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(65424)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12426

Query Match          1.4%; Score 20; DB 4; Length 65424;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1041 GGAGAGAGAGAGAGAAAT 1060
Db      50936 GGAGAGAGAGAGAGAAAT 50917

RESULT 9
US-09-949-016-15868
; Sequence 15868, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15868
; LENGTH: 374159
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15868

Query Match          1.4%; Score 20; DB 4; Length 374159;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      817 TTCTTTCAAAAGATTTCGA 836
Db      265055 TTCTTTCAAAAGATTTCGA 265074

RESULT 10
US-09-949-016-17933/c
; Sequence 17933, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
```

```

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17933
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17933

Query Match          1.4%; Score 19; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      449 TTTAAAAAATTAACCAAG 467
Db      361 TTTAAAAAATTAACCAAG 343

RESULT 11
US-09-949-016-17934/c
; Sequence 17934, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17934
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17934

Query Match          1.4%; Score 19; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      449 TTTAAAAAATTAACCAAG 467
Db      394 TTTAAAAAATTAACCAAG 376

RESULT 12
US-09-949-016-41823/c
; Sequence 41823, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
```

```
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41823
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-41823

Query Match          1.4%; Score 19; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      449 TTTAAATAATACCAAG 467
DB      361 TTTAAATAATACCAAG 343

RESULT 13
US-09-949-016-41824/c
; Sequence 41824, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-09-08
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41824
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-41824

Query Match          1.4%; Score 19; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      449 TTTAAATAATACCAAG 467
DB      394 TTTAAATAATACCAAG 376

RESULT 14
US-09-949-016-145344/c
; Sequence 145344, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
```

```
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 145344
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-145344

Query Match          1.4%; Score 19; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1038 AAGGAGAGAGAGAGAGAG 1056
DB      561 AAGGAGAGAGAGAGAGAG 543

RESULT 15
US-09-949-016-145345/c
; Sequence 145345, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 145345
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-145345

Query Match          1.4%; Score 19; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1038 AAGGAGAGAGAGAGAGAG 1056
DB      553 AAGGAGAGAGAGAGAGAG 535
```

Search completed: September 16, 2005, 08:14:17
Job time : 178.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd

OM nucleic - nucleic search, using SW model

Run on: September 16, 2005, 03:01:45 ; Search time 1291.5 Seconds
(without alignments)
7191.347 Million cell updates/sec

Title:	US-10-713-381-2
Perfect score:	1394
Sequence:	1 ccatggtgtctctatgaaaa.....cttttcgtccaccacatg 1394.

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 7389322 seqs, 333128559 residues

Word size : 0

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Listing first 45 summaries

```
Database : Published Applications NA:*

1: /cgn2_6/ptodataa/1/pubnpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodataa/1/pubnpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodataa/1/pubnpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodataa/1/pubnpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodataa/1/pubnpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodataa/1/pubnpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodataa/1/pubnpna/US06_NEW_PUB.seq.*
8: /cgn2_6/ptodataa/1/pubnpna/US06_PUBCOMB.seq.*
9: /cgn2_6/ptodataa/1/pubnpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodataa/1/pubnpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodataa/1/pubnpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodataa/1/pubnpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodataa/1/pubnpna/US10_PUBCOMB.seq.*
14: /cgn2_6/ptodataa/1/pubnpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodataa/1/pubnpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodataa/1/pubnpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodataa/1/pubnpna/US10F_PUBCOMB.seq.*
18: /cgn2_6/ptodataa/1/pubnpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodataa/1/pubnpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodataa/1/pubnpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodataa/1/pubnpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodataa/1/pubnpna/US10I_NEW_PUB.seq.*
23: /cgn2_6/ptodataa/1/pubnpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodataa/1/pubnpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodataa/1/pubnpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodataa/1/pubnpna/US60_PUBCOMB.seq.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1394	100.0	1394	20	US-10-713-381-2	Sequence 2, Appl1
2	1311	94.0	1394	20	US-10-713-381-1	Sequence 1, Appl1
3	240	17.2	255	20	US-10-713-381-9	Sequence 9, Appl1
4	111	8.0	128	20	US-10-713-381-9	Sequence 3, Appl1
5	65	4.7	1653	20	US-10-733-930-1857	Sequence 1857, Appl1
6	50	3.6	50	20	US-10-713-381-5	Sequence 5, Appl1
7	40	2.9	40	20	US-10-713-381-6	Sequence 6, Appl1

C	9	23	30	2.2	30	20	US-09-923-844-81	Sequence 4, Appl1
C	10	22	1.6	849	9	US-10-085-117-355	Sequence 5, Appl1	
C	11	21	1.5	194945	17	US-10-357-930-56524	Sequence 355, Appl1	
C	12	21	1.5	333	20	US-10-425-115-183040	Sequence 56524, A	
C	13	21	1.5	356	20	US-10-425-115-183040	Sequence 183040, A	
C	14	21	1.5	386	20	US-10-425-115-117183	Sequence 117183, A	
C	15	21	1.5	459	9	US-09-864-761-11465	Sequence 11465, A	
C	16	21	1.5	1203	20	US-10-425-115-51470	Sequence 51470, A	
C	17	21	1.5	1637	18	US-10-425-114-7365	Sequence 7365, A	
C	18	21	1.5	1929	18	US-10-424-599-21199	Sequence 21199, A	
C	19	21	1.5	9265	18	US-10-311-855-2125	Sequence 2125, Appl	
C	20	21	1.5	9265	18	US-10-240-454-49	Sequence 49, Appl	
C	21	21	1.5	18617	15	US-10-240-485-47	Sequence 47, Appl1	
C	22	21	1.5	26493	11	US-09-984-429-240	Sequence 240, Appl	
C	23	20	1.4	255	20	US-10-425-115-114007	Sequence 114007, A	
C	24	20	1.4	257	18	US-10-424-599-139819	Sequence 119819, A	
C	25	20	1.4	260	20	US-10-357-930-60745	Sequence 60745, A	
C	26	20	1.4	311	20	US-10-425-115-176795	Sequence 176795, A	
C	27	20	1.4	605	13	US-10-027-632-139460	Sequence 139460, A	
C	28	20	1.4	605	17	US-10-027-632-139461	Sequence 139461, A	
C	29	20	1.4	605	17	US-10-027-632-139461	Sequence 139461, A	
C	30	20	1.4	680	20	US-10-363-3458-4725	Sequence 4725, Appl	
C	31	20	1.4	680	20	US-10-363-3458-4726	Sequence 4726, Appl	
C	32	20	1.4	680	21	US-10-363-4833-4725	Sequence 4725, Appl	
C	33	20	1.4	680	21	US-10-363-4833-4726	Sequence 4726, Appl	
C	34	20	1.4	821	20	US-10-425-115-34238	Sequence 34238, A	
C	35	20	1.4	973	20	US-10-425-115-63769	Sequence 63769, A	
C	36	20	1.4	1261	20	US-10-425-115-134230	Sequence 134230, A	
C	37	20	1.4	1518	20	US-10-425-115-35658	Sequence 35658, A	
C	38	20	1.4	51198	13	US-10-087-192-1729	Sequence 1729, Appl	
C	39	20	1.4	54016	21	US-10-741-600-17886	Sequence 17886, A	
C	40	20	1.4	72069	21	US-10-741-600-17775	Sequence 17775, A	
C	41	20	1.4	91071	13	US-10-087-192-235	Sequence 235, Appl	
C	42	20	1.4	148497	20	US-10-723-860-1744	Sequence 1744, Appl	
C	43	20	1.4	148497	22	US-10-786-149-675	Sequence 675, Appl	
C	44	20	1.4	337344	19	US-10-388-838-58	Sequence 58, Appl1	
C	45	20	1.4	465337	9	US-09-933-2674-1	Sequence 1, Appl1	

ALIGNMENTS

```

RESULT 1
US-10-713-381-2
; Sequence 2, Application US/10/713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 06/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Zea mays
; US-10-713-381-2

```

Query Match	100.0%	Score 1394	DB 20	Length 1394
Best Local Similarity	100.0%	Pred. No. 0		
Matches 1394	Conservative 0	Mismatches 0	Indels 0	Gaps 0
1	CCATGGTGTCTTATGAAAAGATGATGATCATGTGTCTATTCCTGTTTCTTAGGATCC	60		

Db 1 CCAATGTCCTATGAAAAAGATGAGTCAATGTGTCTATATCCGTTTCTTAGGCTCC 60
Qy 61 CTTCTCTGCTTATTAATGAATCGGGGTTACAAAAAATTTCAAGGGTGCATGAT 120
Db 61 CTTCTCTGCTTATTAATGAATCGGGGTTACAAAAAATTTCAAGGGTGCATGAT 120
Qy 121 CTCATGTTTCACTTCTCCCACTCGGCTGACATTTCTTGGATGCGTGGTCCCAT 180
Db 121 CTCATGTTTCACTTCTCCCACTCGGCTGACATTTCTTGGATGCGTGGTCCCAT 180
Qy 181 CTGACCGAGGCCCATGACACCTTTGAGGACCCATCAAGGGCTTTGAGTGGCCCA 240
Db 181 CTGACCGAGGCCCATGACACCTTTGAGGACCCATCAAGGGCTTTGAGTGGCCCA 240
Qy 241 CGAGAGCTATCCGGTGTGTGTGAATCCAGGGGATATATGTCCTCCCAATGCTCACTATA 300
Db 241 CGAGAGCTATCCGGTGTGTGTGAATCCAGGGGATATATGTCCTCCCAATGCTCACTATA 300
Qy 301 TTATTAATCTTATGATATTAATTAATTTTGGAAAAATTAACAATCTTATCTTTGTGA 360
Db 301 TTATTAATCTTATGATATTAATTAATTTTGGAAAAATTAACAATCTTATCTTTGTGA 360
Qy 361 GGGCTCAGCATGATATTTGCTTAGGGCCAGAAAATGCGAGGACCAAGCATGCTTAGTG 420
Db 361 GGGCTCAGCATGATATTTGCTTAGGGCCAGAAAATGCGAGGACCAAGCATGCTTAGTG 420
Qy 421 TCCACATTTGGCACTTCCAGAACAAAGATTTAAAAAATTAACAAAGTAACTTATCCACT 480
Db 421 TCCACATTTGGCACTTCCAGAACAAAGATTTAAAAAATTAACAAAGTAACTTATCCACT 480
Qy 481 CGAAGCTATCATGATATGTTTAAAGAAACATCTATTTAAACCAAGCATCTTTAAAGAA 540
Db 481 CGAAGCTATCATGATATGTTTAAAGAAACATCTATTTAAACCAAGCATCTTTAAAGAA 540
Qy 541 CAAGCATATTTGCAAGAGACAAATTAATGTTACAGTTTAAACAATCTTAAGCGACAA 600
Db 541 CAAGCATATTTGCAAGAGACAAATTAATGTTACAGTTTAAACAATCTTAAGCGACAA 600
Qy 601 TTATATCGAAAGGTAAGCTATGAGCTTCAATTTCTTTTCATCTTGTATTTTGT 660
Db 601 TTATATCGAAAGGTAAGCTATGAGCTTCAATTTCTTTTCATCTTGTATTTTGT 660
Qy 661 ATGTTTTTATATACATTTTCTTCTTCAATAGAGATTTTCTCCGATTTTATAA 720
Db 661 ATGTTTTTATATACATTTTCTTCTTCAATAGAGATTTTCTCCGATTTTATAA 720
Qy 721 ATGACTATTAAGTCAATTTTATTAAGAGACGACATGCTGATGATTTCTGTTCAAAAATC 780
Db 721 ATGACTATTAAGTCAATTTTATTAAGAGACGACATGCTGATGATTTCTGTTCAAAAATC 780
Qy 781 TTTTGTATTTTAAAGCTAGTTGGCAACCTGTTTCTTCAAGAAATTTGATTT 840
Db 781 TTTTGTATTTTAAAGCTAGTTGGCAACCTGTTTCTTCAAGAAATTTGATTT 840
Qy 841 TTTCAAAAAAATTAATTTTCTTCTTAAATAAGAAAAACATTAAGAAAAATAGAT 900
Db 841 TTTCAAAAAAATTAATTTTCTTCTTAAATAAGAAAAACATTAAGAAAAATAGAT 900
Qy 901 TGGCAGACTAGCCCTAGAAATGTTTCCCAATTAATTAACAATCACTGTGTATTAATTTG 960
Db 901 TGGCAGACTAGCCCTAGAAATGTTTCCCAATTAATTAACAATCACTGTGTATTAATTTG 960
Qy 961 GCGAGCCCAATAATTAATTTAAACGAAATGGAATCGAGGAAACAAATCGAGCTAT 1020
Db 961 GCGAGCCCAATAATTAATTTAAACGAAATGGAATCGAGGAAACAAATCGAGCTAT 1020
Qy 1021 TTTCTAGATTAAGTAAAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Db 1021 TTTCTAGATTAAGTAAAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Qy 1081 TGAATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCTTACGTCA 1140
Db 1081 TGAATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCTTACGTCA 1140

Db 1081 TGAATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCTTACGTCA 1140
Qy 1141 GGTGGGAGGCTCTGCTGATCTCAATGAGATCTACATGCTTTTAAACGTTCCGTC 1200
Db 1141 GGTGGGAGGCTCTGCTGATCTCAATGAGATCTACATGCTTTTAAACGTTCCGTC 1200
Qy 1201 TTGTTTCATTCGTCAAGCCTTCCCTATTTCTGTAACCAAGAGATTAACCTCAAAAT 1260
Db 1201 TTGTTTCATTCGTCAAGCCTTCCCTATTTCTGTAACCAAGAGATTAACCTCAAAAT 1260
Qy 1261 CCATCTTACTCATGCACTTCCATGCAACGACGACATATGTTTCTGTAACGATCTATT 1320
Db 1261 CCATCTTACTCATGCACTTCCATGCAACGACGACATATGTTTCTGTAACGATCTATT 1320
Qy 1321 AAAGATCAACAGGATGAGTTCTCCGCTACTTCCCTCTCCCTGCGCATCTTTT 1380
Db 1321 AAAGATCAACAGGATGAGTTCTCCGCTACTTCCCTCTCCCTGCGCATCTTTT 1380
Qy 1381 CGTCCACGACCATG 1394
Db 1381 CGTCCACGACCATG 1394

RESULT 2
US-10-713-381-1
; Sequence 1, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFEMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: USING SAME
; CURRENT APPLICATION NUMBER: US/10/713.381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-1

Query Match 94.0%; Score 1311; DB 20; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCAATGTCCTATGAAAAAGATGAGTCAATGTGTCTATATCCGTTTCTTAGGCTCC 60
Db 1 CCAATGTCCTATGAAAAAGATGAGTCAATGTGTCTATATCCGTTTCTTAGGCTCC 60
Qy 61 CTTCTCTGCTTATTAATGAATCGGGGTTACAAAAAATTTCAAGGGTGCATGAT 120
Db 61 CTTCTCTGCTTATTAATGAATCGGGGTTACAAAAAATTTCAAGGGTGCATGAT 120
Qy 121 CTCATGTTTCACTTCTCCCACTCGGCTGACATTTCTTGGATGCGTGGTCCCAT 180
Db 121 CTCATGTTTCACTTCTCCCACTCGGCTGACATTTCTTGGATGCGTGGTCCCAT 180
Qy 181 CTGACCGAGGCCCATGACACCTTTGAGGACCCATCAAGGGCTTTGAGTGGCCCA 240
Db 181 CTGACCGAGGCCCATGACACCTTTGAGGACCCATCAAGGGCTTTGAGTGGCCCA 240
Qy 241 CGAGAGCTATCCGGTGTGTGTGAATCCAGGGGATATATGTCCTCCCAATGCTCACTATA 300
Db 241 CGAGAGCTATCCGGTGTGTGTGAATCCAGGGGATATATGTCCTCCCAATGCTCACTATA 300
Qy 301 TTATTAATCTTATGATATTAATTAATTTTGGAAAAATTAACAATCTTATCTTTGTGA 360
Db 301 TTATTAATCTTATGATATTAATTAATTTTGGAAAAATTAACAATCTTATCTTTGTGA 360

```

Db 301 TTTATTTCTTTAGATTTATTTAATTTTGGAAAAATACAACTTATCTTTGGTGA 360
Qy 361 GGGGCTCAGCATGATTTTGGCTTAGGGCCAGAAATGCGAGAACCGACCTGTCTAGTG 420
Db 361 GGGGCTCAGCATGATTTTGGCTTAGGGCCAGAAATGCGAGAACCGACCTGTCTAGTG 420
Qy 421 TCCACTATTGGGCTACCCGAGAACCAAGATTTAAAAAATAACCAAGTAACTTAATCCACT 480
Db 421 TCCACTATTGGGCTACCCGAGAACCAAGATTTAAAAAATAACCAAGTAACTTAATCCACT 480
Qy 481 CGAAGCTATCATGATATGTTTAAAGAACATCTAATTAACCAAGATCTCTTAAAAAA 540
Db 481 CGAAGCTATCATGATATGTTTAAAGAACATCTAATTAACCAAGATCTCTTAAAAAA 540
Qy 541 CAAGCATATTTGCAAAAGACAAATATGTTTACAGTTTACAAACATCTTAAGCGACAAA 600
Db 541 CAAGCATATTTGCAAAAGACAAATATGTTTACAGTTTACAAACATCTTAAGCGACAAA 600
Qy 601 TTTATATCGAAAGGTAAGCTATGACGTTCAAGATTTTCTTTTCAATCTTGTTATTTTGT 660
Db 601 TTTATATCGAAAGGTAAGCTATGACGTTCAAGATTTTCTTTTCAATCTTGTTATTTTGT 660
Qy 661 ATTGTTTTTATATACATTTTCTTCTTCAATAGAGTATTTCTTCCGATTTTATAA 720
Db 661 ATTGTTTTTATATACATTTTCTTCTTCAATAGAGTATTTCTTCCGATTTTATAA 720
Qy 721 ATGACTATTAAGCATTTTATATTAAGAGACGACATGTCGTAGATTTCTGTTCAAAAATC 780
Db 721 ATGACTATTAAGCATTTTATATTAAGAGACGACATGTCGTAGATTTCTGTTCAAAAATC 780
Qy 781 TTTCTGATTTTATTAAGAGTATTTGCGAAACCCCTGTTTCTTCAAGAAATTTGATTT 840
Db 781 TTTCTGATTTTATTAAGAGTATTTGCGAAACCCCTGTTTCTTCAAGAAATTTGATTT 840
Qy 841 TTTCAAAAAAATAGTTTATTTTCTCTTATTAATAAGAAACACTTGAATAATAGAGT 900
Db 841 TTTCAAAAAAATAGTTTATTTTCTCTTATTAATAAGAAACACTTGAATAATAGAGT 900
Qy 901 TGGCAGACTAGCCCTGTAAGATGTTTCCCAATAAATTAACAATCATCTGTGTATTAATTTG 960
Db 901 TGGCAGACTAGCCCTGTAAGATGTTTCCCAATAAATTAACAATCATCTGTGTATTAATTTG 960
Qy 961 GCGAGCCCATTAATTTATTTAAACGAAATCGAAATCGAGCCAAATCTGAGCTAT 1020
Db 961 GCGAGCCCATTAATTTATTTAAACGAAATCGAAATCGAGCCAAATCTGAGCTAT 1020
Qy 1021 TTTCTAGATTAAGTAAAAAGGAGAGAGAGAGAGAGAAATCAGTTTAAAGTCAATGCTCC 1080
Db 1021 TTTCTAGATTAAGTAAAAAGGAGAGAGAGAGAGAGAAATCAGTTTAAAGTCAATGCTCC 1080
Qy 1081 TGAAGATGTCGGTTTGGCAACGATAGCACCGTAAATCATAGCTCATAGGTGCTTACGTCA 1140
Db 1081 TGAAGATGTCGGTTTGGCAACGATAGCACCGTAAATCATAGGTGCTTACGTCA 1140
Qy 1141 GGTTCGGGAGCTCTCGTGCATCTCAATGCGCATTAATAGTGTTCACACCGTTCGTC 1200
Db 1141 GGTTCGGGAGCTCTCGTGCATCTCAATGCGCATTAATAGTGTTCACACCGTTCGTC 1200
Qy 1201 TTGTTTCATCGTCCAGACCTTGCTTATTTGAAACCAAGAGATTACTTCCCAACAT 1260
Db 1201 TTGTTTCATCGTCCAGACCTTGCTTATTTGAAACCAAGAGATTACTTCCCAACAT 1260
Qy 1261 CCATCTTACTCATGCAACTTCCATGCAAAACGCGACATATGTTTCTGAAC 1311
Db 1261 CCATCTTACTCATGCAACTTCCATGCAAAACGCGACATATGTTTCTGAAC 1311

```

RESULT 3
 US-10-713-381-9
 ; Sequence 9, Application US/10713381
 ; Publication No. US20040221331A1
 ; GENERAL INFORMATION:

```

; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 255
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-9

Query Match      17.2%; Score 240; DB 20; Length 255;
Best Local Similarity 100.0%; Pred. No. 5.8e-110;
Matches 240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1155 CGTGTCAATGCACATGCGCATCTACATGCTTGTCAACCGTTCGTTGTCATGCTCC 1214
Db 15 CGTGTCAATGCACATGCGCATCTACATGCTTGTCAACCGTTCGTTGTCATGCTCC 74
Qy 1215 AAGCTTGGCTATTTCGAACCAAGAGATACCTTCCCAACAAATCCATCTTACTCATG 1274
Db 75 AAGCTTGGCTATTTCGAACCAAGAGATACCTTCCCAACAAATCCATCTTACTCATG 134
Qy 1275 CAATTCATGCAAAACGCGACATATGTTTCTGTAACAGATATTAAGATCAACAAG 1334
Db 135 CAATTCATGCAAAACGCGACATATGTTTCTGTAACAGATATTAAGATCAACAAG 194
Qy 1335 CTAGGCTTCCCGGCTAGCTTCCCTCCCTCCGATCTTTTTCGCCACACAG 1394
Db 195 CTAGGCTTCCCGGCTAGCTTCCCTCCCTCCGATCTTTTTCGCCACACAG 254

```

RESULT 4
 US-10-713-381-3
 ; Sequence 3, Application US/10713381
 ; Publication No. US20040221331A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ALBERTSEN, MARC C.
 ; APPLICANT: FOX, TIMOTHY W.
 ; APPLICANT: GARNAT, CARL W.
 ; APPLICANT: HUFFMAN, GARY
 ; APPLICANT: KENDALL, TIMMY L.
 ; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
 ; FILE REFERENCE: 578R
 ; CURRENT APPLICATION NUMBER: US/10/713,381
 ; PRIOR FILING DATE: 2003-11-14
 ; PRIOR APPLICATION NUMBER: 08/880,499
 ; NUMBER OF SEQ ID NOS: 24
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 3
 ; LENGTH: 158
 ; TYPE: DNA
 ; ORGANISM: Zea mays

Query Match 8.0%; Score 111; DB 20; Length 158;
 Best Local Similarity 100.0%; Pred. No. 6.2e-45;
 Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1201 TTGTTTCATCGTCCAGACCTTGCTTATTTGAAACCAAGAGATTACTTCCCAACAT 1260
Db 48 TTGTTTCATCGTCCAGACCTTGCTTATTTGAAACCAAGAGATTACTTCCCAACAT 107

```

Oy 1261 CCATCTTACATGCAACTTCATCGAACAAGCACAATATGTTCCGTGAAC 1311
|||
Db 108 CCACTTACTCATGCAACTTCATCGAACAAGCACAATATGTTCCGTGAAC 158

```

RESULT 5
US-10-739-930-1857
; Sequence 1857, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53377) B
; CURRENT APPLICATION NUMBER: US/10/739, 930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 1857
; LENGTH: 1663
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: ZEAMA-23APR03-CLUSTER1429_8
US-10-739-930-1857

```

```

Query Match      4.7%; Score 65; DB 20; Length 1663;
Best Local Similarity 100.0%; Pred. No. 1.2e-21;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1318 ATTTAAGATCACAA CAGCTGACGCTTCTCCCGTAGCTTCCCTCTCTCTGCGCATCTT 13777

Db 6 ATTTAAGATCACAA CAGCTGACGCTTCTCCCGTAGCTTCCCTCTCTCTGCGCATCTT 65

QY	1378	TTTTCG	1382
Db	66	TTTTCG	70

```

RESULT 6
US-10-713-381-5
; Sequence 5, Application US/10713381
; Publication No. US2004022131A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARRAAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY J.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: USING SAME
; FILE REFERENCE: 579R
; CURRENT APPLICATION NUMBER: US/10-713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-5

```

Query Match	3.6%;	Score 50;	DB 20;	Length 50;
Best Local Similarity	100.0%;	Pred. No. 3.3e-14;		
Matches 50;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

[illegible]

```

RESULT 7
US-10-713-381-6
; Sequence 6, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARMAAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-6

```

Query Match 2.9%; Score 40; DB 20; Length 40;
Best Local Similarity 100.0%; Pred. No. 3.5e-09;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

	QY	1239	AGGATACCTACTCCCAACAATCATCTTACTGCATCGAAC	1278
Dδ		1	AGGATACTTAATCTCCAAACAATCATTCTTACTGCATCGAAC	40

```

RESULT 8
US-10-713-381-4
; Sequence 4, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: NALe TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-4

```

Query Match 2.2%; Score 30; DB 20; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.00038;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1179	CATGCTGTCAACCGTTCGTTGTCCA	1208
Db	1	CATGCTGTCAACCGTTCGTTGTCCA	30

RESULT 9
US-09-923-84B-5/C
; Sequence 5, Application US/09923844B
; Patent No. US20020166143A1
; GENERAL INFORMATION:

```

; APPLICANT: Pioneer Hi-Bred International, Inc.
; APPLICANT: Bao, Zhongmeng
; TITLE OF INVENTION: Sclerotinia-inducible Genes and
; FILE REFERENCE: 35718/234631
; CURRENT APPLICATION NUMBER: US/09/923,844B
; PRIOR FILING DATE: 2001-08-07
; PRIOR APPLICATION NUMBER: US 60/224,603
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 849
; TYPE: DNA
; ORGANISM: Helianthus annuus
; US-09-923-844B-5

Query Match          1.6%; Score 23; DB 9; Length 849;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      854 AGTTATTTCTCTTTATAAAT 876
Db      556 AGTTATTTCTCTTTATAAAT 534

RESULT 10
US-10-085-117-355/C
; Sequence 355, Application US/10085117
; Publication No. US2003023334A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: 529452000121
; CURRENT APPLICATION NUMBER: US/10/085,117
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 355
; LENGTH: 194945
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: variation
; LOCATION: (1)...(194945)
; OTHER INFORMATION: n = any nucleotide
; US-10-085-117-355

Query Match          1.6%; Score 22; DB 17; Length 194945;
Best Local Similarity 100.0%; Pred. No. 8.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      113 TGCATGATCTCATGTTTCACT 134
Db      28972 TGCATGATCTCATGTTTCACT 28951

RESULT 11
US-10-357-930-56524
; Sequence 56524, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; FILE REFERENCE: HUMAN PROSTATE CANCER
; APPLICANT: MRI-007BCN
```

```

; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 56524
; LENGTH: 333
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-357-930-56524

Query Match          1.5%; Score 21; DB 20; Length 333;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      93 TTACAAAAAAGTTCCAGGGT 113
Db      89 TTACAAAAAAGTTCCAGGGT 109

RESULT 12
US-10-425-115-183040
; Sequence 183040, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 183040
; LENGTH: 356
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_98517C.1
; US-10-425-115-183040

Query Match          1.5%; Score 21; DB 20; Length 356;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      632 TTTTCTTTTTCATTTCTTGT 652
Db      72 TTTTCTTTTTCATTTCTTGT 92

RESULT 13
US-10-425-115-117183/C
; Sequence 117183, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
```



```
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(53222)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 117183
/ LENGTH: 386
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MRT4577_3835C.1
US-10-425-115-117183

Query Match          1.5%; Score 21; DB 20; Length 386;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      371 ATGATTTTCGCTTAGGGCCC 391
Db      367 ATGATTTTCGCTTAGGGCCC 347

RESULT 14
US-09-864-761-11465/C
/ Sequence 11465, Application US/09864761
/ Patent No. US20020048763A1
/ GENERAL INFORMATION:
/ APPLICANT: Penn, Sharron G.
/ APPLICANT: Rank, David R.
/ APPLICANT: Hanzel, David K.
/ APPLICANT: Chen, Wensheng
/ TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
/ FILE REFERENCE: Aecm1ca-X-1
/ CURRENT APPLICATION NUMBER: US/09/864,761
/ CURRENT FILING DATE: 2001-05-23
/ PRIOR APPLICATION NUMBER: US 60/180,312
/ PRIOR FILING DATE: 2000-02-04
/ PRIOR APPLICATION NUMBER: US 60/207,456
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: US 09/632,366
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR APPLICATION NUMBER: GB 24263.6
/ PRIOR FILING DATE: 2000-10-04
/ PRIOR APPLICATION NUMBER: US 60/236,359
/ PRIOR FILING DATE: 2000-09-27
/ PRIOR APPLICATION NUMBER: PCT/US01/00666
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00667
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00664
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00669
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00665
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00668
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00663
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00662
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00661
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00670
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: US 60/234,687
/ PRIOR FILING DATE: 2000-09-21
/ PRIOR APPLICATION NUMBER: US 09/608,408
/ PRIOR FILING DATE: 2000-06-30
/ PRIOR APPLICATION NUMBER: US 09/774,203
```

```
/ PRIOR FILING DATE: 2001-01-29
/ NUMBER OF SEQ ID NOS: 49117
/ SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
/ SEQ ID NO 11465
/ LENGTH: 459
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO AC000100.2
/ OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.59
/ OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.59
/ OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.59
/ OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.6
/ OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.64
/ OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.2
US-09-864-761-11465

Query Match          1.5%; Score 21; DB 9; Length 459;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1037 AAAGGAGAGAGAGAGAGA 1057
Db      118 AAAGGAGAGAGAGAGAGA 98

RESULT 15
US-10-425-115-51470/C
/ Sequence 51470, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(53222)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 51470
/ LENGTH: 1203
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (1)..(1203)
/ OTHER INFORMATION: unsure at all n locations
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MRT4577_146936C.1
US-10-425-115-51470

Query Match          1.5%; Score 21; DB 20; Length 1203;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      848 AAATTAGTTATTTCTCTT 868
Db      1129 AAATTAGTTATTTCTCTT 1109
```

Search completed: September 16, 2005, 12:49:11
Job time : 1293.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2005, 10:24:35 ; Search time 4293.5 Seconds
(without alignments)
12358.587 Million cell updates/sec

Title: US-10-713-381-2

Perfect score: 1394
Sequence: 1 cccatggtcgtcctcatgaaaa.....cttttcgtccaccaccatg 1394

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsa1:*
9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	908	65.1	963	9	CC656933 OGMDO20TM
2	667	47.8	915	9	CG224225 OGIAG08TV
3	409	29.3	687	9	CC656939 OGMDO20TV
4	66	4.7	878	9	CG224211 OGIAG08TV
5	27	1.9	539	8	B2828689 B2828689
6	27	1.9	611	8	B2828694 B2828694
7	27	1.9	804	9	CG148339 PUFH74TD
8	26	1.9	854	9	CG011564 ZUBA73TV
9	25	1.8	687	9	B2371637 B2371637
10	25	1.8	726	9	CC810403 ZMMBC047
11	25	1.8	819	9	CG305519 OGMDO20TV
12	25	1.8	825	9	CG190606 PUIH82TB
13	25	1.8	833	9	CG190609 PUIH82TB
14	25	1.8	857	8	CC358144 PCHP12TD
15	25	1.8	878	9	CG307825 OGMBC01TH
16	25	1.8	885	8	CG307825 OGMBC01TH
17	25	1.8	892	9	CG197610 PUPW51TD
18	25	1.8	893	8	CC439146 PUIH56TB
19	25	1.8	956	9	CG060573 PUCN71TD
20	24	1.7	385	1	A1626382 A1626382
21	24	1.7	468	1	A1626141 A1626141
22	24	1.7	502	8	BH411902 BH411902
23	24	1.7	533	8	CC010354 CC010354
24	24	1.7	625	9	CG377165 CG377165

25	24	1.7	630	9	CG331717 OGCXK74TH
26	24	1.7	699	9	CM009494 ZMMBLA001
27	24	1.7	739	9	CM009496 ZMMBLA001
28	24	1.7	835	9	CC979335 ZUACSA47TV
29	24	1.7	857	8	CC335124 OGCXK74TH
30	24	1.7	866	9	CG209882 OGIAB01TH
31	23	1.6	229	8	CC364841 PUIHBD06TD
32	23	1.6	318	9	CG421210 ZMMBC003
33	23	1.6	493	1	AL375883 ZMMBLA001
34	23	1.6	514	8	B2032120 B2032120
35	23	1.6	533	8	AZ095114 RPCT-23-1
36	23	1.6	610	4	B1676219 B1676219
37	23	1.6	651	8	BH297118 BH297118
38	23	1.6	664	9	CG700282 ZMMBC011
39	23	1.6	691	8	B2524893 OGAQ075TC
40	23	1.6	710	8	B2828501 B2828501
41	23	1.6	831	8	B2129958 CH230-503
42	23	1.6	833	9	CG437195 OGVHC05TV
43	23	1.6	845	9	CG255896 OGYAT70TV
44	23	1.6	870	9	CL198382 ZMMBC025
45	23	1.6	900	9	CT708844 OGMCG92TV

ALIGNMENTS

RESULT 1	CC656933	963 bp	DNA	linear	GSS 19-JUN-2003
LOCUS	OGMDO20TM_ZM_0.7_1.5_KB				
DEFINITION	genomic survey sequence.				
ACCESSION	CC656933				
VERSION	CC656933.1	GI:32060225			
KEYWORDS	GSS.				
SOURCE	Zea mays				
ORGANISM	Zea mays				
REFERENCE	Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utecht, T., Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea. 1 (bases 1 to 963)				
AUTHORS	Whitelaw, C.A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.				
TITLE	Unpublished (2002)				
JOURNAL	Consortium for Maize Genomics				
COMMENT	Other GSSs: OGMDO20TV Contact: Cathy Whitelaw TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA Tel: 301-838-5843 Fax: 301-838-0208 Email: whitelaw@cigr.org Seq primer: TR Class: sheared ends.				
FEATURES	Location/Qualifiers				
source	1..963 /organism="Zea mays" /mol_type="genomic DNA" /strain="B73" /db_xref="taxon:4577" /clone="ZMMMA0554D15" /clone_1ib="ZM_0.7_1.5_KB" /note="Vector: pBCSK-1 site 1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"				
ORIGIN					
Query Match	65.1%;	Score 908;	DB 9;	Length 963;	
Best Local Similarity	100.0%;	Pred. No. 0;			
Matches 908;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	404	ACACGACATGTCTAGTGTCCATATTCGACATCCAGCAAGATTTAAAAATTAACC	463		
DB	1	ACCACGATGTCTAGTGTCCATATTCGACATCCAGCAAGATTTAAAAATTAACC	60		

```
QY 464 AAGTAAGTATCTACGCGAAGCTATCATGTATGTTTAAAGAAACATCTATTAAACC 523
Db 61 AAGTAAGTATCTACGCGAAGCTATCATGTATGTTTAAAGAAACATCTATTAAACC 120
QY 524 AGCATCCTCTTAAACAAAGCATATTTGCAAGAGCAAAATATGTTACAGTTTACAA 583
Db 121 AGCATCCTCTTAAACAAAGCATATTTGCAAGAGCAAAATATGTTACAGTTTACAA 180
QY 584 CATCTAAGGCGCAAAATATATCGAAGGTAGCTATGACGTTGAGATTTTCTTTTC 643
Db 181 CATCTAAGGCGCAAAATATATCGAAGGTAGCTATGACGTTGAGATTTTCTTTTC 240
QY 644 ATTCCTGTTATTTGTTATGTTTATATACATTTTCTCTTACAAATAGAGTATT 703
Db 241 ATTCCTGTTATTTGTTATGTTTATATACATTTTCTCTTACAAATAGAGTATT 300
QY 704 TCTTCGATTTTAAATGACTATATAAGTCAATTTTATATAAGACGACGATGTGAG 763
Db 301 TCTTCGATTTTAAATGACTATATAAGTCAATTTTATATAAGACGACGATGTGAG 360
QY 764 ATTCGTGTTCAAAATCTTCTGATTTTATAGACTAGTTGGCAACCTGTCTTT 823
Db 361 ATTCGTGTTCAAAATCTTCTGATTTTATAGACTAGTTGGCAACCTGTCTTT 420
QY 824 CAAGAATTTGATTTTCAAAAAAATAGTTTATTTTCTCTTATATAAGAAAC 883
Db 421 CAAGAATTTGATTTTCAAAAAAATAGTTTATTTTCTCTTATATAAGAAAC 480
QY 884 ACTTAGAAAAATAGAGTGGCAGACTAGCCCTAGAAATGTTTCCCAATAATTAACA 943
Db 481 ACTTAGAAAAATAGAGTGGCAGACTAGCCCTAGAAATGTTTCCCAATAATTAACA 540
QY 944 CTGTGATATATTTTGGCAGGCCCATTAATTTTAAACGGAACGTAATCGAGCA 1003
Db 541 CTGTGATATATTTTGGCAGGCCCATTAATTTTAAACGGAACGTAATCGAGCA 600
QY 1004 AACCAATCTGAGCTATTTCTCTAGATTAGTAAAGAGGAGAGAGAGAGAAATCAG 1063
Db 601 AACCAATCTGAGCTATTTCTCTAGATTAGTAAAGAGGAGAGAGAGAAATCAG 660
QY 1064 TTTTAAGTCAATGCTCCCTAGATGTGGCGGTTGGCAACGATGACCAAGTATCAT 1123
Db 661 TTTTAAGTCAATGCTCCCTAGATGTGGCGGTTGGCAACGATGATCATAGCT 720
QY 1124 CATAGTGCTAGCTAGGTTGGCAGCTCTGCTCATCTCAAGGCAATCTATATCAG 1183
Db 721 CATAGTGCTAGCTAGGTTGGCAGCTCTGCTCATCTCAAGGCAATCTATATCAG 780
QY 1184 TTGTTCAACGTTGCTGTTGTTTCATGTCGAAGCCTTGCTATTTCTGAACCAAGAGAT 1243
Db 781 TTGTTCAACGTTGCTGTTGTTTCATGTCGAAGCCTTGCTATTTCTGAACCAAGAGAT 840
QY 1244 ACCTACTCCCAAAACATTCATCTTACTCATGCAACTTCAGAAACGACATATGTT 1303
Db 841 ACCTACTCCCAAAACATTCATCTTACTCATGCAACTTCAGAAACGACATATGTT 900
QY 1304 TCCTGAAC 1311
Db 901 TCCTGAAC 908

RESULT 2
CG224225 915 bp DNA linear GSS 22-AUG-2003
DEFINITION OGIAG08TV_ZM_0.7_1.5_KB Zea mays genomic clone ZMMBma071B51,
genomic survey sequence.
ACCESSION CG224225
VERSION CG224225.1 GI:34124113
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```

```
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
OIGIAG08TV
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@cigr.org
Seq primer: TP
Classes: sheared ends.

FEATURES
source
1..915
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBma071B15"
/clone_1ib="ZM 0.7 1.5 KB"
/note="Vector: pBCK-1; Site 1: HindII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match 47.8%; Score 667; DB 9; Length 915;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 667; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 645 TTCTGTTATTTGTTATGTTTATATACATTTCTCTCTTCAATAGAGTATT 704
Db 24 TTCTGTTATTTGTTATGTTTATATACATTTCTCTCTTCAATAGAGTATT 83
QY 705 CTTCGATTTTAAATAGACTATAAGTCAATTTTATATAAGACGACGATGTGAGA 764
Db 84 CTTCGATTTTAAATAGACTATAAGTCAATTTTATATAAGACGACGATGTGAGA 143
QY 765 TTCTGTTCAAAATCTTCTGATTTTAAAGCTAGTTGGCAACCTGTCTTTC 824
Db 144 TTCTGTTCAAAATCTTCTGATTTTAAAGCTAGTTGGCAACCTGTCTTTC 203
QY 825 AAAGAATTTGATTTTCAAAAAAATAGTTTATTTCTCTTATATAAGAAACA 884
Db 204 AAAGAATTTGATTTTCAAAAAAATAGTTTATTTCTCTTATATAAGAAACA 263
QY 885 CTTAGAAAAATAGAGTGGCAGACTAGCCCTAGAAATGTTTCCCAATAATTAACA 944
Db 264 CTTAGAAAAATAGAGTGGCAGACTAGCCCTAGAAATGTTTCCCAATAATTAACA 323
QY 945 TGTGATATATTTTGGCCAGGCCCATTAATTTTAAACGGAACGTAATCGAGCA 1004
Db 324 TGTGATATATTTTGGCCAGGCCCATTAATTTTAAACGGAACGTAATCGAGCA 383
QY 1005 ACCAATCTGAGCTATTTCTCTAGATTAGTAAAGGAGAGAGAGAGAAATTCAGT 1064
Db 384 ACCAATCTGAGCTATTTCTCTAGATTAGTAAAGGAGAGAGAGAGAAATTCAGT 443
QY 1065 TTTAAGTCAATGCTCCCTAGATGTGGCGGTTGGCAACGATGACCAAGTATCATAGCT 1124
Db 444 TTTAAGTCAATGCTCCCTAGATGTGGCGGTTGGCAACGATGACCAAGTATCATAGCT 503
QY 1125 ATAGTGCTTACGTCAGGTTGGCAGCTCTGCTGATCTCAATGCGCATCTACATGCT 1184
Db 504 ATAGTGCTTACGTCAGGTTGGCAGCTCTGCTGATCTCAATGCGCATCTACATGCT 563
QY 1185 TGTTCACCGTTGCTGTTGTTTCATGTCGAAGCCTTGCTATTTCTGAACCAAGAGATA 1244
Db 564 TGTTCACCGTTGCTGTTGTTTCATGTCGAAGCCTTGCTATTTCTGAACCAAGAGATA 623
```

Qy	1203	GTTCACATGTCACAGCCTTGCTCTTTCTTGAACCAAGAGGATCTACTCTCCAAACATCC	1262
Db	373	GTTCACATGTCACAGCCTTGCTCTTTCTTGAACCAAGAGGATCTACTCTCCAAACATCC	314
Qy	1263	ATCTTACATGACAACTTCATGCAACACGCAATATGTTCTCTGAC	1311
Db	313	ATCTTACATGACAACTTCATGCAACACGCAATATGTTCTCTGAC	265
RESULT 4			
LOCUS	CG224211/c	878 bp	DNA
DEFINITION	OG1AG08TH_ZM_0.7_1.5_KB_Zea_mays genomic clone ZMMBHa0716B15,	linear	GSS 22-AUG-2003
ACCESSION	CG224211		
VERSION	CG224211.1		GI:34124099
KEYWORDS	GSS.		
SOURCE			
ORGANISM	Zea mays		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.		
REFERENCE	1 (bases 1 to 878)		
AUTHORS	WhiteJaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Reenick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nunez,A., Robbins,D. and Lakey,N.		
TITLE	Consortium for Maize Genomics		
JOURNAL	Unpublished (2002)		
COMMENT	Other GSSs: OG1AG08TV		
	Contact: Cathy WhiteJaw		
	TIGR		
	9712 Medical Center Drive, Rockville, MD 20850, USA		
	Tel: 301-838-5843		
	Fax: 301-838-0208		
	Email: whiteJaw@tigr.org		
	Seq primer: TR		
	Class: sheared ends.		
FEATURES			
Source	Location/Qualifiers		
	1..878		
	/organism="Zea mays"		
	/mol_type="genomic DNA"		
	/strain="B73"		
	/db_xref="taxon:4577"		
	/clone="ZMMBHa0716B15"		
	/clone_1bp="ZM_0.7_1.5_KB"		
	/note="Vector: pBCKS-; Site 1: HincII; 0.7-1.5 kb		
	methylation filtered genomic DNA library"		
ORIGIN			
Query Match	4.7%;	Score 66;	DB 9;
Best Local Similarity	100.0%;	Pred. No. 5.5e-22;	Length 878;
Matches	66;	Conservative 0;	Mismatches 0;
		Indels 0;	Gaps 0;
Qy	1324	GATCACAACAGCTGCGCTCCCGCTAGGCTCCCTCTCTCTGCGAGTCTTTTTCGT	1383
Db	878	GATCACAACAGCTGCGCTCCCGCTAGGCTCCCTCTCTCTGCGAGTCTTTTTCGT	819
Qy	1384	CCACCA 1389	
Db	818	CCACCA 813	
RESULT 5			
LOCUS	B2828689/c	539 bp	DNA
DEFINITION	PURHW4TB_ZM_0.6_1.0_KB_Zea_mays genomic clone ZMMBHa320N03,	linear	GSS 18-MAR-2003
ACCESSION	B2828689		
VERSION	B2828689.1		GI:29050496
KEYWORDS	GSS.		
SOURCE			
ORGANISM	Zea mays		

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 539)
WhiteJaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T., Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and Bennettzen,J.
Maize Genomics Consortium
Unpublished (2003)
Other GSSs: PUFHM74TD
Contact: Cathy WhiteJaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteJaw@tigr.org
Seq primer: TR
Classes: sheared ends.

FEATURES
source
1..539
/organism="Zea mays"
/mol_type="genomic DNA"
/db_xref="B73"
/db_xref="taxon:4577"
/clone_lib="ZM 0.6 1.0 KB"
/note="Vector: PCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
COT selected genomic DNA library"

ORIGIN
Query Match 1.9%; Score 27; DB 8; Length 539;
Best Local Similarity 100.0%; Pred. No. 0.052;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 843 CAAAAAAATTAGTTATTTCTCTTT 869
|||||
292 CAAAAAAATTAGTTATTTCTCTTT 266

RESULT 6
LOCUS BZ828694 611 bp DNA linear GSS 18-MAR-2003
DEFINITION PUFHM74TD ZM 0.6 1.0 KB Zea mays genomic clone ZM8BTA320N03,
genomic survey sequence.
ACCESSION BZ828694
VERSION BZ828694.1 GI:29050508
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 611)
WhiteJaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T., Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and Bennettzen,J.
Maize Genomics Consortium
Unpublished (2003)
Other GSSs: PUFHM74TD
Contact: Cathy WhiteJaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteJaw@tigr.org
Seq primer: TR
Classes: sheared ends.

FEATURES
source
1..611
/organism="Zea mays"
/mol_type="genomic DNA"
/db_xref="B73"
/db_xref="taxon:4577"
/clone_lib="ZM8BTA320N03"

/clone_lib="ZM 0.6 1.0 KB"
/note="Vector: PCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
COT selected genomic DNA library"

ORIGIN
Query Match 1.9%; Score 27; DB 8; Length 611;
Best Local Similarity 100.0%; Pred. No. 0.052;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 843 CAAAAAAATTAGTTATTTCTCTTT 869
|||||
Db 222 CAAAAAAATTAGTTATTTCTCTTT 248

RESULT 7
LOCUS CG148339 804 bp DNA linear GSS 21-AUG-2003
DEFINITION PUKB646TB ZM 0.6 1.0 KB Zea mays genomic clone ZM8BTA0780G19,
genomic survey sequence.
ACCESSION CG148339
VERSION CG148339.1 GI:34039122
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 804)
WhiteJaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T., Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and Bennettzen,J.
Maize Genomics Consortium
Unpublished (2003)
Other GSSs: PUKB646TD
Contact: Cathy WhiteJaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteJaw@tigr.org
Seq primer: TR
Classes: sheared ends.

FEATURES
source
1..804
/organism="Zea mays"
/mol_type="genomic DNA"
/db_xref="B73"
/db_xref="taxon:4577"
/clone_lib="ZM8BTA0780G19"
/note="Vector: PCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
COT selected genomic DNA library"

ORIGIN
Query Match 1.9%; Score 27; DB 9; Length 804;
Best Local Similarity 100.0%; Pred. No. 0.052;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 837 TTTTTCAAAAAAATTAGTTATTTT 863
|||||
Db 156 TTTTTCAAAAAAATTAGTTATTTT 182

RESULT 8
LOCUS CG011564/c 854 bp DNA linear GSS 19-AUG-2003
DEFINITION ZUABT23TV ZM 3.0 4.0 KB Zea mays genomic clone ZM8BPA0018C22,
genomic survey sequence.
ACCESSION CG011564
VERSION CG011564.1 GI:33883730
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 854)

TITLE
JOURNAL
COMMENT
Other_GSSS: Z048T23TH
Contact: Cathy Whitelaw
TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.

FEATURES
source
location/Qualifiers

1..854
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM58Pa0018C22"
/clone_1lb="ZM_3.0.4.0_KB"
/note="Vector: pBSCsk-; Site_1: HincII; 3-4 kb 'unfiltered' genomic DNA library"

ORIGIN

Query Match
Best Local Similarity 100.0%; Pred. No. 0.17; Length 854;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 268 GGGGATATATGTCTCCACATATGTC 293
Db 221 GGGGATATATGTCTCCACATATGTC 196

RESULT 9
BZ371637/c 687 bp DNA linear GSS 26-NOV-2002
LOCUS 1e38n03.5', genomic survey sequence.
DEFINITION
BZ371637
BZ371637
BZ371637.1 GI:25455537

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 687)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Unpublished (2002)
Contact: W. Richard McCombie
Lita Annenberg Hazen Genome Sequencing Center
Cold Spring Harbor Laboratory
PO Box 100 Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8884
Fax: 516 367 8874
Email: mccombie@cshl.org
Place: 1e38 row: h column: 03
Seq primer: -21M13Unlvfwd
Class: shotgun
High quality sequence stop: 687.
Location/Qualifiers

FEATURES
source
1..687
/organism="Zea mays"
/mol_type="genomic DNA"

ORIGIN

Query Match
Best Local Similarity 100.0%; Pred. No. 0.55; Length 687;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 845 AAAAAAATGATTATTTCTCTT 869
Db 363 AAAAAAATGATTATTTCTCTT 339

RESULT 10
CC810403/c 726 bp DNA linear GSS 16-JUL-2003
LOCUS ZM58Bc0479C04r ZM58Bc Zea mays genomic clone ZM58Bc0479C04 3',
DEFINITION
BZ371637/c 726 bp DNA linear GSS 16-JUL-2003
LOCUS 1e38n03.5', genomic survey sequence.
BZ371637
BZ371637.1 GI:32817028

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 726)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Unpublished (2003)
Contact: Bharti, A.K.
Dr. Joachim Messing's lab
The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers University
190 Frelinghuysen Road, Piscataway, NJ 08854, USA
Tel: 732 445 3801
Fax: 732 445 5735
Email: bharti@waksman.rutgers.edu
Seq primer: SP6
Class: BAC ends
High quality sequence start: 397.
Location/Qualifiers

FEATURES
source
1..726
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ZM58Bc0479C04"
/lab_host="E. coli DH10B"
/clone_1lb="ZM58Bc"
/note="Vector: pTARBAC1.3; Site_1: BamHI; Site_2: BamHI"

ORIGIN

Query Match
Best Local Similarity 100.0%; Pred. No. 0.55; Length 726;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 265 CCAGGGATATATGTCTCCACAT 289
Db 364 CCAGGGATATATGTCTCCACAT 340

```

RESULT 11
CG305519/c      819 bp      DNA      linear      GSS 25-AUG-2003
LOCUS           CG305519
DEFINITION      CG305519 ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMA0713C15,
                 genomic survey sequence.
ACCESSION       CG305519
VERSION         CG305519.1
KEYWORDS        GSS.
SOURCE          Zea mays
ORGANISM        Zea mays
REFERENCE       Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
                 clade; Panicoidae; Andropogoneae; Zea.
                 1 (bases 1 to 819)
                 Whitelew,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
                 Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
                 Clack,R.W., Nuneberg,A., Robbins,D. and Lakey,N.
                 Consortium for Maize Genomics
                 Unpublished (2002)
                 Other GSSs: CG305519
                 Contact: Cathy Whitelew
TIGR            9712 Medical Center Drive, Rockville, MD 20850, USA
                 Tel: 301-838-5843
                 Fax: 301-838-0208
                 Email: whitelew@tigr.org
                 Seq primer: TP
                 Class: sheared ends.
FEATURES
source          Location/Qualifiers
                 1..819
                 /organism="Zea mays"
                 /mol_type="genomic DNA"
                 /strain="B73"
                 /db_xref="taxon:4577"
                 /clone_1lb="ZM_0.7_1.5_KB"
                 /note="Vector: pBCSK-; Site 1: HindII; 0.7-1.5 kb
                 methylation filtered genomic DNA library"
ORIGIN
Query Match      1.8%; Score 25; DB 9; Length 819;
Best Local Similarity 100.0%; Pred. No. 0.55;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 844 AAAAAAATTAGTTATTTCTCTT 868
Db 142 AAAAAAATTAGTTATTTCTCTT 118

```

```

Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelew@tigr.org
Seq primer: TP
Class: sheared ends.
FEATURES
source          Location/Qualifiers
                 1..825
                 /organism="Zea mays"
                 /mol_type="genomic DNA"
                 /strain="B73"
                 /db_xref="taxon:4577"
                 /clone_1lb="ZMMBTA0585M20"
                 /note="Vector: pCR4-TOP0; Site 1: EcoRI; 0.6-1.0 kb high
                 cot selected genomic DNA library"
ORIGIN
Query Match      1.8%; Score 25; DB 9; Length 825;
Best Local Similarity 100.0%; Pred. No. 0.55;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 845 AAAAAAATTAGTTATTTCTCTT 869
Db 715 AAAAAAATTAGTTATTTCTCTT 691

```



```

9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: white1aw@ci9r.org
Seq primer: TR
Class: sheared ends.
FEATURES
    location/Qualifiers
        source
            1..878
                /organism="Zea mays"
                /mol_type="genomic DNA"
                /strain="B73"
                /db_xref="taxon:4577"
                /clone="ZWMEMa0527A01"
                /clone_1lb="ZM_0.7.1.5_KB"
                /note="Vector: pBCSK-; Site_1.HincII; 0.7-1.5 kb
                methylation filtered genomic DNA library"
ORIGIN
Query Match          1.8%; Score 25; DB 9; Length 878;
Best Local Similarity 100.0%; Pred.No.0.55;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      844  AAAAAAAAAATTAGTTATTTTCTCTT      868
        |||||||
        167  AAAAAAAAAATTAGTTATTTTCTCTT      143
Search completed: September 16, 2005, 17:48:56
Job time : 4293.5 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 08:25:31 ; Search time 4677.65 Seconds

(without alignments)
14440.280 Million cell updates/sec

Title: US-10-713-381-2

Perfect score: 1394
Sequence: 1 ccattggtgctctcatgaaac.....cttttcgctccaccaccatg 1394

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sta:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1394	100.0	1394	6	AX224395 Sequence
2	1394	100.0	1394	6	BD062177 Male clis
3	1389.2	99.7	1394	6	AX224394 Sequence
4	1389.2	99.7	1394	6	BD062176 Male clis
5	1387.6	99.5	1394	6	AF360356 Zea mays
6	243.6	17.5	255	6	AX224402 Sequence
7	146	10.5	158	6	AX224396 Sequence
8	77	5.5	186199	2	AC147602 Zea mays
9	73.4	5.3	25769	3	AC117267 Dictyoste
10	71.2	5.1	713	3	AF034389 Plasmodi
11	71.2	5.1	3576	3	DDU86962 Dictyoste
12	70.8	5.1	149526	2	DX957346 Danio rer
13	70.4	5.1	347582	3	PFMAL4P1 Plasmodi
14	70.2	5.0	883	6	CQ422641 Sequence
15	70	5.0	241	11	BV119878 PZA01377
16	70	5.0	247	11	BV119874 PZA01377
17	70	5.0	248	11	BV119877 PZA01377
18	69.8	5.0	13400	6	CO85052 Sequence
19	69.6	5.0	139583	2	AC151050 Zea mays

C	20	69.4	5.0	110000	8	CR382132_17	Continuation (18 C
	21	69.4	5.0	181381	2	CR450818	CR450818 Danio rer
	22	69.2	5.0	114817	2	AC149836	AC149836 Zea mays
	23	69.2	5.0	266544	3	AC116956	AC116956 Dictyoste
	24	69	4.9	164443	2	CR751221	CR751221 Danio rer
	25	68.8	4.9	7347	1	AF211124	AF211124 Carionell
	26	68.6	4.9	186199	2	AC147602	AC147602 Zea mays
	27	68.4	4.9	213	11	BV119875	BV119875 PZA01377
	28	68.4	4.9	216	11	BV119871	BV119871 PZA01377
	29	68.4	4.9	224	11	BV119882	BV119882 PZA01377
	30	68.4	4.9	245	11	BV119881	BV119881 PZA01377
	31	68.4	4.9	246	11	BV119869	BV119869 PZA01377
	32	68.2	4.9	346296	8	AF090447	AF090447 Zea mays
	33	68	4.9	110000	2	PFMAL7P1_02	Continuation (3 of
	34	67.8	4.9	227495	2	AC148480	AC148480 Zea mays
	35	67.6	4.8	351	11	BV110530	BV110530 PZA01059
	36	67.6	4.8	331039	3	AC116979	AC116979 Dictyoste
	37	67.4	4.8	170	11	BV119873	BV119873 PZA01377
	38	67	4.8	2781	3	AY044085	AY044085 Dictyoste
	39	67	4.8	110000	3	AC116984_1	Continuation (2 of
	40	66.8	4.8	228	11	BV119870	BV119870 PZA01377
	41	66.8	4.8	234	11	BV119883	AF338825 Cloning v
	42	66.8	4.8	7774	12	AF338825	AF338825 Cloning v
	43	66.8	4.8	8658	12	AF338824	AF338824 Cloning v
	44	66.8	4.8	12029	3	AE001434	AE001434 Plasmodi
	45	66.8	4.8	47573	3	AF030694	AF030694 Plasmodi

ALIGNMENTS

RESULT 1	AX224395	Sequence 2 from Patent WO0160997.	1394 bp	DNA	1linear	PAT 10-SEP-2001
LOCUS	AX224395					
DEFINITION	Sequence 2 from Patent WO0160997.					
ACCESSION	AX224395					
VERSION	AX224395.1	GI:15554637				
KEYWORDS						
SOURCE						
ORGANISM	Zea mays					
	Zea mays					
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.					
REFERENCE	1	Albertsen, M.C., Fox, T.W., Garnaat, C.W., Huffman, G. and Kendall, T.L.				
AUTHORS						
TITLE	Male tissue-preferred regulatory region and method of using same					
JOURNAL	Patent: WO 0160997-A 2 23-AUG-2001.					
	PIONEER HI-BRED INTERNATIONAL, INC. (US)					
FEATURES	Location/Qualifiers					
source	1. .1394					
	/organism="Zea mays"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:4577"					
ORIGIN						
Query Match	100.0%; Score 1394; DB 6; Length 1394;					
Best Local Similarity	100.0%; Pred. No. 4.1e-269;					
Matches 1394; Conservative	0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	CCATGCTGCTCTGAGAAAAGATGATGATGCTGATGCTGCTTTCTTAGGCTCC	60			
DB	1	CCATGCTGCTCTGAGAAAAGATGATGATGCTGATGCTGCTTTCTTAGGCTCC	60			
QY	61	CTTCTCTGCTTATTAAGTACGATCGAGGTTACAAAACCTTCACGAGTGCATGAT	120			
DB	61	CTTCTCTGCTTATTAAGTACGATCGAGGTTACAAAACCTTCACGAGTGCATGAT	120			
QY	121	CTTCATGTTCCATCTTCCACCTCGCGTTGACATTTCTTGATGTCGTTGCCAT	180			
DB	121	CTTCATGTTCCATCTTCCACCTCGCGTTGACATTTCTTGATGTCGTTGCCAT	180			
QY	181	CTGACGAGGCCCATAGACACCTTCCGGAACCCATCAAGGCTTCGATGCCCA	240			
DB	181	CTGACGAGGCCCATAGACACCTTCCGGAACCCATCAAGGCTTCGATGCCCA	240			

Db 181 CTGACGAGGCCCATGACACACCTTTCGGAGACCCATCAAGGAGCCTTTCGATGCGCA 240
QY 241 CGAGAGTATCGGGTGTGGTATCCAGGGATATATGTCGCCCAAGATGTCACCTATA 300
Db 241 CGAGAGTATCGGGTGTGGTATCCAGGGATATATGTCGCCCAAGATGTCACCTATA 300
QY 301 TTATATTTCTTTAGATATATTTAAATTTTGGAAAAATACAAATCTTATCTTTGTGA 360
Db 301 TTATATTTCTTTAGATATATTTAAATTTTGGAAAAATACAAATCTTATCTTTGTGA 360
QY 361 GGGCTTCAGATGATTTTTCGTTAGGGCCAGAAATCGAGAGACCAAGCATCTAGTG 420
Db 361 GGGCTTCAGATGATTTTTCGTTAGGGCCAGAAATCGAGAGACCAAGCATCTAGTG 420
QY 421 TCCACTATTGGACTACCCAGAACAGATTTTAAATAATACCAAGTATCTATCCACT 480
Db 421 TCCACTATTGGACTACCCAGAACAGATTTTAAATAATACCAAGTATCTATCCACT 480
QY 481 CGAAGCTATCATGTATGTTTAAAGAAACATCTATTTAAACACGATCCTTTAAAAA 540
Db 481 CGAAGCTATCATGTATGTTTAAAGAAACATCTATTTAAACACGATCCTTTAAAAA 540
QY 541 CAAGCATTTTGCAGAGACCAATTTATGTTACAGTTTACAAATCTTAAAGCCGACAA 600
Db 541 CAAGCATTTTGCAGAGACCAATTTATGTTACAGTTTACAAATCTTAAAGCCGACAA 600
QY 601 TTATATCGAAGGTAAGCTATGACGTTACAGATTTTCTTTTCATCTTGTATTTTGT 660
Db 601 TTATATCGAAGGTAAGCTATGACGTTACAGATTTTCTTTTCATCTTGTATTTTGT 660
QY 661 ATTTGTTTTATATACATTTTCTTCTTACAAATAGATGATTTTCTCCGATTTTAAAA 720
Db 661 ATTTGTTTTATATACATTTTCTTCTTACAAATAGATGATTTTCTCCGATTTTAAAA 720
QY 721 ATGACATATAAGCATTTTATATAAGAGACCATGTCGTATCTGTTCAAAAATC 780
Db 721 ATGACATATAAGCATTTTATATAAGAGACCATGTCGTATCTGTTCAAAAATC 780
QY 781 TTTCTGATTTTAAAGCTAGTTTGGCAACCTGTTTCTTCAAGAATTTTGAATTT 840
Db 781 TTTCTGATTTTAAAGCTAGTTTGGCAACCTGTTTCTTCAAGAATTTTGAATTT 840
QY 841 TTTCAAAAAAATAGTTTATTTTCTCTTATATAAATAGAAAAACATTTAGAAAAATAG 900
Db 841 TTTCAAAAAAATAGTTTATTTTCTCTTATATAAATAGAAAAACATTTAGAAAAATAG 900
QY 901 TGGCAACTAGCCCTAGAAATGTTTCCCAATAATTAACAATCACTGTGTATTTTGG 960
Db 901 TGGCAACTAGCCCTAGAAATGTTTCCCAATAATTAACAATCACTGTGTATTTTGG 960
QY 961 GCCAGCCCAATAAATATTAAACGAAACTGAATCGAGCGAAACCAATCTGAGCTAT 1020
Db 961 GCCAGCCCAATAAATATTAAACGAAACTGAATCGAGCGAAACCAATCTGAGCTAT 1020
QY 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAGAAATCACTTTTAAATGCTATG 1080
Db 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAGAAATCACTTTTAAATGCTATG 1080
QY 1081 TGAAGATGTCGGTTTGGCAACGATAGCCATCCGTAATCATAGCTCATAGTGTCTAG 1140
Db 1081 TGAAGATGTCGGTTTGGCAACGATAGCCATCCGTAATCATAGTGTCTAGTGA 1140
QY 1141 GGTTCGGAGCTCTCGTGTATCTCATATGAGCATCTACATGCTTGTAAACGTTTCG 1200
Db 1141 GGTTCGGAGCTCTCGTGTATCTCATATGAGCATCTACATGCTTGTAAACGTTTCG 1200
QY 1201 TTGTTCCATGTCGAAGCCTTGGCTATTTGAAACGAAGATACCTACTCCCAACAT 1260
Db 1201 TTGTTCCATGTCGAAGCCTTGGCTATTTGAAACGAAGATACCTACTCCCAACAT 1260
QY 1261 CCATCTTACATGCAATTCATGCAACAGGCAATATGTTTCTGTAACGATCTATT 1320
Db 1261 CCATCTTACATGCAATTCATGCAACAGGCAATATGTTTCTGTAACGATCTATT 1320

QY 1321 AAAGATCACAAACAGCTAGGTTCTCCGCTAGCTTCCCTCTGCTGCGGATCTTTT 1380
Db 1321 AAAGATCACAAACAGCTAGGTTCTCCGCTAGCTTCCCTCTGCTGCGGATCTTTT 1380
QY 1381 CGTCCAGCACCATG 1394
Db 1381 CGTCCAGCACCATG 1394
RESULT 2
BD062177 1394 bp DNA linear PAT 27-AUG-2002
LOCUS BD062177
DEFINITION Male tissue-preferred regulatory region and method of using same.
ACCESSION BD062177
VERSION BD062177.1 GI:22607782
KEYWORDS JP 2001520523-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1394) Fox,T.W., Garmae,C.W., Huffman,G.A. and
Kendall,T.L.
AUTHORS Male tissue-preferred regulatory region and method of using same
TITLE PATENT: JP 2001520523-A 2 30-OCT-2001;
JOURNAL PIONEER HI BRED INTERNATIONAL, INC
COMMENT PN JP 2001520523-A/2
PD 30-OCT-2001
PF 19-JUN-1998 JP 1999504910
PR 23-JUN-1997 US 08/880499
PI MARC C ALBERTSEN,TIMOTHY W FOX,CARL W GARMAE,GARY A HUFFMAN,
PI TIMMY L KENDALL
PC C12N15/82,C12N15/29,C12N9/24,C12N9/22,C12N9/10,C12N9/00 PC
,C07K14/34,C12Q1/68,
PC A01H5/00
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers.
FEATURES
source Location/Qualifiers
1..1394
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 1394; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 4,1e-269;
Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCATGCTGCTCTATGAAAAAGATGATGATGATGATGATGATGATGATGATGATGATG 60
Db 1 CCATGCTGCTCTATGAAAAAGATGATGATGATGATGATGATGATGATGATGATGATG 60
QY 61 CTTCTTTCGCTTATTAATGATCTGAATCGGGGTTACAAAAAATTTCAAGGGTGCATGAT 120
Db 61 CTTCTTTCGCTTATTAATGATCTGAATCGGGGTTACAAAAAATTTCAAGGGTGCATGAT 120
QY 121 CTCATGTTCCACTTCTCCACCTGCGGTGACATTTCTTGGAGTGGGTGTTCCCAT 180
Db 121 CTCATGTTCCACTTCTCCACCTGCGGTGACATTTCTTGGAGTGGGTGTTCCCAT 180
QY 181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCA 240
Db 181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCA 240
QY 241 CGAGAGTATCGGGTGTGGTATCCAGGGATATATGTCGCCCAAGATGTCACCTATA 300
Db 241 CGAGAGTATCGGGTGTGGTATCCAGGGATATATGTCGCCCAAGATGTCACCTATA 300
QY 301 TTATATTTCTTTAGATATATTTTAAATTTTGGAAAAATACAAATCTTATCTTTGTGA 360
Db 301 TTATATTTCTTTAGATATATTTTAAATTTTGGAAAAATACAAATCTTATCTTTGTGA 360

Oy	361	GGGCTCAGCAATGATTTTTCGTTAGGGGCCGAATATCGAGGACACGACATGCTATAGT	420
Db	361	GGGCTCAGCAATGATTTTTCGTTAGGGGCCGAATATCGAGGACACGACATGCTATAGT	420
Oy	421	TCGACTATTTGGCACTACCCAGAACAGATTTTAAAAAAATTAACCAAGTAACTAATCCACT	480
Db	421	TCGACTATTTGGCACTACCCAGAACAGATTTTAAAAAAATTAACCAAGTAACTAATCCACT	480
Oy	481	CGAAAGCTATCATGTAAATGTTTAAAGAAACATCTATTAAAAACGAGATCCCTTTAAAAA	540
Db	481	CGAAAGCTATCATGTAAATGTTTAAAGAAACATCTATTAAAAACGAGATCCCTTTAAAAA	540
Oy	541	CAAGCATATTTTCGAAGAGAGACAAATTTATGTTACAGTTTACAAACATCTTAAGGCGACAA	600
Db	541	CAAGCATATTTTCGAAGAGAGACAAATTTATGTTACAGTTTACAAACATCTTAAGGCGACAA	600
Oy	601	TTATATCGAAAGGTACGATGACGTTGAGATTTTCTTTTCATTCCTGTATATTTGTT	660
Db	601	TTATATCGAAAGGTACGATGACGTTGAGATTTTCTTTTCATTCCTGTATATTTGTT	660
Oy	661	ATTGTTTTTAAATACATTTTCTTCCTTAACAATAGAGATTTTCTTCGATTTTATATAA	720
Db	661	ATTGTTTTTAAATACATTTTCTTCCTTAACAATAGAGATTTTCTTCGATTTTATATAA	720
Oy	721	ATGACTATTAAGTCAATTTTATATATAGACAGCATGTCTGATATTCCTGTTCAAAATC	780
Db	721	ATGACTATTAAGTCAATTTTATATATAGACAGCATGTCTGATATTCCTGTTCAAAATC	780
Oy	781	TTTTCGATTTTTTTTAAAGCTAGTTGGCAACCCGTGTTCTTTCAAAAGATTTGATTTT	840
Db	781	TTTTCGATTTTTTTTAAAGCTAGTTGGCAACCCGTGTTCTTTCAAAAGATTTGATTTT	840
Oy	841	TTCAAAAAAAATTAAGTTATTTCTCTTTATAAATAGAAAAACCTTAGAAAAATAGAGT	900
Db	841	TTCAAAAAAAATTAAGTTATTTCTCTTTATAAATAGAAAAACCTTAGAAAAATAGAGT	900
Oy	901	TGCCAGACTAGCCCTAGATGTTTTCCCAATTAATTAACAATCACTGTGTATATTAATTTG	960
Db	901	TGCCAGACTAGCCCTAGATGTTTTCCCAATTAATTAACAATCACTGTGTATATTAATTTG	960
Oy	961	GCGAGCCCCCATTAATTAATTTAAACCGGAATCGAATTCGAGGAAACCAATCTGAGCAT	1020
Db	961	GCGAGCCCCCATTAATTAATTTAAACCGGAATCGAATTCGAGGAAACCAATCTGAGCAT	1020
Oy	1021	TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAGAAATCAGTTTAACTATGTCATGTC	1080
Db	1021	TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAGAAATCAGTTTAACTATGTCATGTC	1080
Oy	1081	TGAGATGTGCGGTTTGGCAACGATAGCCAGTAACTCATAGCTACCTACGTC	1140
Db	1081	TGAGATGTGCGGTTTGGCAACGATAGCCAGTAACTCATAGCTACCTACGTC	1140
Oy	1141	GSTTGGGAGGCTCTGTGTCTCATCTCACATGGCATCTCAATGCTTTGTTCAACCGTTGCTC	1200
Db	1141	GSTTGGGAGGCTCTGTGTCTCATCTCACATGGCATCTCAATGCTTTGTTCAACCGTTGCTC	1200
Oy	1201	TTGTTTCATGCTCGAAGCCTTGCTCATTTCTGAACCAAGAGATACCTACCCAAACAT	1260
Db	1201	TTGTTTCATGCTCGAAGCCTTGCTCATTTCTGAACCAAGAGATACCTACCCAAACAT	1260
Oy	1261	CCATCTTACTATGCAACTTCCATGCAAAACGACACATATGTTCCTGAACAGATCTATTT	1320
Db	1261	CCATCTTACTATGCAACTTCCATGCAAAACGACACATATGTTCCTGAACAGATCTATTT	1320
Oy	1321	AAAGATCACAACAGCTAGCGTTCTTCGCTAGCTTCCCTCTCTCTCGCCGATCTTTTT	1380
Db	1321	AAAGATCACAACAGCTAGCGTTCTTCGCTAGCTTCCCTCTCTCTCGCCGATCTTTTT	1380
Oy	1381	CGTCACCAACCATG 1394	
Db	1381	CGTCACCAACCATG 1394	

LOCUS	AX224394	1394 bp	DNA	linear	PAT 10-SEP-2001
DEFINITION	Sequence 1 from Patent WO0160997.				
ACCESSION	AX224394				
VERSION	AX224394.1	GI:15554636			
KEYWORDS					
SOURCE					
ORGANISM	Zea mays				
REFERENCE	1				
AUTHORS	Albertsen, M. C., Fox, T. W., Garnaat, C. W., Huffman, G. and Kendall, T. L.				
TITLE	Male tissue-preferred regulatory region and method of using same				
JOURNAL	Patent: WO 0160997-A 1 23-AUG-2001;				
FEATURES	PIONEER HI-BRED INTERNATIONAL, INC. (US)				
source	Location/Qualifiers				
ORIGIN	1..1394				
	/organism="Zea mays"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:4577"				
Query Match	99.7%; Score 1389.2; DB 6; Length 1394;				
Best Local Similarity	99.8%; Pred. No. 3.8e-268;				
Matches 1391; Conservative	0; Mismatches 3; Indels 0; Gaps 0;				
OY	1	CCATGCTGTCTCTATGAAAAAGATGATGACAAATGCTCTATATCCGTTTCTTAAAGGTTCC	60		
DB	1	CCATGGTGCTCTATGAAAAAGATGATGACAAATGCTCTATATCCGTTTCTTAAAGGTTCC	60		
OY	61	CTTCTTCTGCTTATTAATCTGACTGTAATCGGAGGTTACAAAACCTTCCACGGGTGCATGAT	120		
DB	61	CTTCTTCTGCTTATTAATCTGACTGTAATCGGAGGTTACAAAACCTTCCACGGGTGCATGAT	120		
OY	121	CTGCATGTTCCAGTCTTCCCACTTCGCGCTTGACATTTCTTGAATGTCGATGTTCCCAT	180		
DB	121	CTGCATGTTCCAGTCTTCCCACTTCGCGCTTGACATTTCTTGAATGTCGATGTTCCCAT	180		
OY	181	CTGACCGAGGCGCCATCGACACCTTTGGGACACCCATGAAGGGCTTTCCGATGGGCCCA	240		
DB	181	CTGACCGAGGCGCCATCGACACCTTTGGGACACCCATGAAGGGCTTTCCGATGGGCCCA	240		
OY	241	CGAACAGTATGGGGTCTGATGATCCAGGGGATATATGTCGCCCAATCGTCACTTATA	300		
DB	241	CGAACAGTATGGGGTCTGATGATCCAGGGGATATATGTCGCCCAATCGTCACTTATA	300		
OY	301	TTATTAATCTTAAATTAATTAATTAATTTTGAATAATTAACAACTTAATACCTTTTGTTGA	360		
DB	301	TTATTAATCTTAAATTAATTAATTAATTTTGAATAATTAACAACTTAATACCTTTTGTTGA	360		
OY	361	GGGCTCGACGATATATTTGGCTTAAGGGCCCGAAGAAAGCGAGGACCGAGCCATGCTATAG	420		
DB	361	GGGCTCGACGATATATTTGGCTTAAGGGCCCGAAGAAAGCGAGGACCGAGCCATGCTATAG	420		
OY	421	TCGACTATGGCACTACCCAGAACAAAGTTTAAAAAATAACCAAGTAACTAATCCACT	480		
DB	421	TCGACTATGGCACTACCCAGAACAAAGTTTAAAAAATAACCAAGTAACTAATCCACT	480		
OY	481	CGAAGCTATCATGTATATGTTTAAAGAAACATCTAATTAACCAAGTAACTAATCCACT	540		
DB	481	CGAAGCTATCATGTATATGTTTAAAGAAACATCTAATTAACCAAGTAACTAATCCACT	540		
OY	541	CAAGCATATTTGAAAGAGCAAAATTAATGTTACAGTTTAAACCAATCTTAAGAGGACAAA	600		
DB	541	CAAGCATATTTGAAAGAGCAAAATTAATGTTACAGTTTAAACCAATCTTAAGAGGACAAA	600		
OY	601	TTAATATGAAAGGTAAGCTATGACGTTCAAGTTTCTTTTCAATCTTGTTATTTGTT	660		
DB	601	TTAATATGAAAGGTAAGCTATGACGTTCAAGTTTCTTTTCAATCTTGTTATTTGTT	660		

Db	781	TTTTGATTTTTTTTAAAGACTAGTTGGCACCCCTGTTCTTCAAAGAAATTTTGATTTT	840
Oy	841	TTCAAAAAAATTAGTTTATTTCTCTTTATMAATAGAAAAACCTTAGAAAAATAGAGT	900
Db	841	TTCAAAAAAATTAGTTTATTTCTCTTTATMAATAGAAAAACCTTAGAAAAATAGAGT	900
Oy	901	TGCCAGACTAGCCCTTAGAATGTTTTCCCATMAATTACATCAGTGTGTATTAATTATGG	960
Db	901	TGCCAGACTAGCCCTTAGAATGTTTTCCCATMAATTACATCAGTGTGTATTAATTATGG	960
Oy	961	GCGAGCCCATTAATTATTTAAACCGAAATCGAATTCGAGGAAACCAATCTAGCAT	1020
Db	961	GCGAGCCCATTAATTATTTAAACCGAAATCGAATTCGAGGAAACCAATCTAGCAT	1020
Oy	1021	TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAATCAGTTTAACTATGTTCC	1080
Db	1021	TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAATCAGTTTAACTATGTTCC	1080
Oy	1081	TGAGATGTCGGTGTGGCAACGATAGCCACCGTAATCATAGCTCATAGCTCTAGCTCA	1140
Db	1081	TGAGATGTCGGTGTGGCAACGATAGCCACCGTAATCATAGCTCATAGCTCTAGCTCA	1140
Oy	1141	GTTTCGAGAGCTCTGTGTCACTTCATAGCATAGCATAGCTGTGTTCACACCGTTCGTC	1200
Db	1141	GTTTCGAGAGCTCTGTGTCACTTCATAGCATAGCATAGCTGTGTTCACACCGTTCGTC	1200
Oy	1201	TTGTTCCATCGTCCAAAGCCTTGCTTATCTGAAACCAAGAGATACCTACTCCAAACAT	1260
Db	1201	TTGTTCCATCGTCCAAAGCCTTGCTTATCTGAAACCAAGAGATACCTACTCCAAACAT	1260
Oy	1261	CCATCTTACTCATGCAACTTCCATGCAAAACGCAATATGTTTCTGTAACAGATCTATT	1320
Db	1261	CCATCTTACTCATGCAACTTCCATGCAAAACGCAATATGTTTCTGTAACAGATCTATT	1320
Oy	1321	AAAGATCAACAAGCTAGCGTTTCTCCCGTAGCTTCCCTCTCTCTCTGCGCATCTTTT	1380
Db	1321	AAAGATCAACAAGCTAGCGTTTCTCCCGTAGCTTCCCTCTCTCTCTGCGCATCTTTT	1380
Oy	1381	CGTCACACCAACATG 1394	
Db	1381	CGTCACACCAACATG 1394	
RESULT 5	AF360356	3343 bp DNA linear	PLN 12-MAY-2001
LOCUS	AF360356		
DEFINITION	Zea mays male fertility protein (Ms45) gene, complete cds.		
VERSION	AF360356.1	GI:14028756	
KEYWORDS			
SOURCE			
ORGANISM	Zea mays		
REFERENCE	Bukaryova, V.Iridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoidae; Andropogoneae; Zea.		
AUTHORS	1 (bases 1 to 3343)		
TITLE	Fox,T.W., Trimnell,M.R. and Albertsen,M.C.		
JOURNAL	Cloning of Ms45, a gene required for male fertility from Zea mays		
REFERENCE	Unpublished		
AUTHORS	2 (bases 1 to 3343)		
TITLE	Fox,T.W., Trimnell,M.R. and Albertsen,M.C.		
JOURNAL	Direct Submision		
REFERENCE	Submitted (13-MAR-2001) Trait and Technology Development, Pioneer		
AUTHORS	Hi-Bred Intl. Inc., 7300 N.W. 62nd Ave., P.O. Box 1004, Johnston,		
TITLE	IA 50131-1004, USA		
JOURNAL			
FEATURES			
SOURCE	Location/Qualifiers		
	1..3343		
	/organism="Zea mays"		
	/mol type="genomic DNA"		
	/culivar="B73"		
	/db_xref="taxon:4577"		
	/chromosome="9L"		

gene	<1392..->2942
mrna	/gene="M845" join(<1392..1768,1898..2182,2280..2447,2534..->2942) /gene="M845" /product="male fertility protein" join(1392..1768,1898..2182,2280..2447,2534..2942) /gene="M845"
CDS	/codon_start=1 /product="male fertility protein" /protein_id="AAK52489.1" /db_xref="GI:14028757" /translation="MEKRNLOMRGRGDIQVPHLPALALALVADPGLSLPELV DYPVYKELAPYGEWMSMPRDNASRLRRLRLEFVGVEPSEIEDLOGRGYAGLA DGRVVRWGEAGMETFAVNPWMSCEANGVSTKROKHEEBFCGRPLGRLTFEIG TGBELVADVAYGLMVGQSGVAVSVAERADGPIRANLDVHRGSVPFTIRSMY SRKHNLILLEGEGTGLRLYRDETSQSHVVLGLVFPNQVLSIEHQFLFSETTCK RIMKYLEGRAGEVEFANLPGPDPVRSNGQGFVVALDCCRTPAQVEFARPMLR TLYKFPRLSKVLTWKARRMHYVALLDSEGRVVELERGRHEVKKLVSEVREVRK LWIGTVAHNHIAITIPYLED"
Query Match	99.5%; Score 1387.6; DB 8; Length 3343;
Best Local Similarity	99.7%; Pred. No. 77e-268;
Matches 1390; Conservative	0; Mismatches 4; Indels 0; Gaps 0;
Db	1 CCATGGTGTCTTAGAAAAAGATGATGACATGTGTCATATATCCGTTTTCTTAGGGTCC 60
Oy	1 CCATGGTGTCTTAGAAAAAGATGATGACATGTGTCATATATCCGTTTTCTTAGGGTCC 60
Db	1 CCATGGTGTCTTAGAAAAAGATGATGACATGTGTCATATATCCGTTTTCTTAGGGTCC 60
Oy	61 CTTCTTTCGCTTATTAAGTGAATCGGGGTTACAAAACTTCCAGGGTGCATGAT 120
Db	61 CTTCTTTCGCTTATTAAGTGAATCGGGGTTACAAAACTTCCAGGGTGCATGAT 120
Oy	121 CTCCATGTTCCATCTTCCACCGCGGTGACATCTTGGATGTGGGTGTTCCCAT 180
Db	121 CTCCATGTTCCATCTTCCACCGCGGTGACATCTTGGATGTGGGTGTTCCCAT 180
Oy	181 CTGACCGAGGCCATCAGACACCTTTGGGAGCACCCATCAAGGGCTTTGGATGGCCA 240
Db	181 CTGACCGAGGCCATCAGACACCTTTGGGAGCACCCATCAAGGGCTTTGGATGGCCA 240
Oy	241 CGAAGCGTATCGGGTCTGTGTATCCAGGGGATATATGTCCCCCAATGTGACCTATA 300
Db	241 CGAAGCGTATCGGGTCTGTGTATCCAGGGGATATATGTCCCCCAATGTGACCTATA 300
Oy	301 TTATTTATCTTAGATATTTTAAATTTTGGAAAAATTAACAACCTATCTTTGTGTA 360
Db	301 TTATTTATCTTAGATATTTTAAATTTTGGAAAAATTAACAACCTATCTTTGTGTA 360
Oy	361 GGGCTCAGCATAGATTTTCGTTAGGGCCAGAAATGCGAGNACCGACCATGTCTAGTG 420
Db	361 GGGCTCAGCATAGATTTTCGTTAGGGCCAGAAATGCGAGNACCGACCATGTCTAGTG 420
Oy	421 TCACGATATGGCACTACCCAGAACAGATTTAAAAAAATTAACCAAGTAACTAATCCACT 480
Db	421 TCACGATATGGCACTACCCAGAACAGATTTAAAAAAATTAACCAAGTAACTAATCCACT 480
Oy	481 CGAAGCATATCATATATGTTTAAAGAAACATCTAATTAACAACGATCCTCTTAABAAA 540
Db	481 CGAAGCATATCATATATGTTTAAAGAAACATCTAATTAACAACGATCCTCTTAABAAA 540
Oy	541 CAAAGCATATTCGAAAGAGCAAAATTAATGTTAAGTTTAAACCAACGATCCTTAABAAA 600
Db	541 CAAAGCATATTCGAAAGAGCAAAATTAATGTTAAGTTTAAACCAACGATCCTTAABAAA 600
Oy	601 TTATATGAAAGTAGCTATGACGTTCCAGATTTTCTTTTCATCTTGTTATTTTGT 660
Db	601 TTATATGAAAGTAGCTATGACGTTCCAGATTTTCTTTTCATCTTGTTATTTTGT 660
Oy	661 ATTGTTTTTATATACATTTCTCTCTTAATATATGATGATTTTCTTCCGATTTTATATA 720
Db	661 ATTGTTTTTATATACATTTCTCTCTTAATATATGATGATTTTCTTCCGATTTTATATA 720

QY 721 ATGACTATAAAGTCATTTTATATTAAGACGACGATGCTAGATTCTCGTCAAAAATC 780
 |||||||
 DB 721 ATGACTATAAAGTCATTTTATATTAAGACGACGATGCTAGATTCTCGTCAAAAATC 780
 |||||||
 QY 781 TTTCTGATTTTTTAAAGCTAGTTGGCAACCTGTTCTTTCAAAAATTTTGAATTT 840
 |||||||
 DB 781 TTTCTGATTTTTTAAAGCTAGTTGGCAACCTGTTCTTTCAAAAATTTTGAATTT 840
 |||||||
 QY 841 TTCAAAAAATAGTTTATTTTCTTTTAAATAAGAAAACACTAGAAAAATAGACT 900
 |||||||
 DB 841 TTCAAAAAATAGTTTATTTTCTTTTAAATAAGAAAACACTAGAAAAATAGACT 900
 |||||||
 QY 901 TGCCAGACTAGCCCTAGAAATGTTTCCCATTAATTAACAATCACTGTATTAATTTTG 960
 |||||||
 DB 901 TGCCAGACTAGCCCTAGAAATGTTTCCCATTAATTAACAATCACTGTATTAATTTTG 960
 |||||||
 QY 961 GCCAGCCCCATAAATATTATTTAAACGAAACTGAATCGAGCAACCAATCTGAGCTAT 1020
 |||||||
 DB 961 GCCAGCCCCATAAATATTATTTAAACGAAACTGAATCGAGCAACCAATCTGAGCTAT 1020
 |||||||
 QY 1021 TTCTCTAGANTTACTAAAAAGGAGAGAGAGAGAAATCACTTTTAAGTCAATGCTCC 1080
 |||||||
 DB 1021 TTCTCTAGANTTACTAAAAAGGAGAGAGAGAGAAATCACTTTTAAGTCAATGCTCC 1080
 |||||||
 QY 1081 TGAGATGTGCGGTTTGGCAAGATAGCAACCGTAATCATAGCTCATAGGTGCTTACGTCA 1140
 |||||||
 DB 1081 TGAGATGTGCGGTTTGGCAAGATAGCAACCGTAATCATAGCTCATAGGTGCTTACGTCA 1140
 |||||||
 QY 1141 GGTTCGGCAGCTCTCGTGTATTTCAACATGCACTACTAGCTGTTGTTCAACCGTTCGTC 1200
 |||||||
 DB 1141 GGTTCGGCAGCTCTCGTGTATTTCAACATGCACTACTAGCTGTTGTTCAACCGTTCGTC 1200
 |||||||
 QY 1201 TTGTTCCATGTCGCAAGCCTTGCTATTTGAACCAAGAGATTAATCTCCCAACAT 1260
 |||||||
 DB 1201 TTGTTCCATGTCGCAAGCCTTGCTATTTGAACCAAGAGATTAATCTCCCAACAT 1260
 |||||||
 QY 1261 CCATCTTACTCATGCAACTTCCATGCAACCGCAGATATGTTTCTTGAACAGATCTAT 1320
 |||||||
 DB 1261 CCATCTTACTCATGCAACTTCCATGCAACCGCAGATATGTTTCTTGAACAGATCTAT 1320
 |||||||
 QY 1321 AAGATTCACACAGCTAGGCTTTCGCCGTAAGCTTCTCTCTCTGCGCATCTTTT 1380
 |||||||
 DB 1321 AAGATTCACACAGCTAGGCTTTCGCCGTAAGCTTCTCTCTCTGCGCATCTTTT 1380
 |||||||
 QY 1381 CGTCCACGACCATG 1394
 |||||||
 DB 1381 CGTCCACGACCATG 1394
 |||||||

RESULT 6
 AX224402 255 bp DNA 1linear PAT 10-SEP-2001
 LOCUS AX224402
 DEFINITION Sequence 9 from Patent WO0160997.
 ACCESSION AX224402
 VERSION AX224402.1 GI:15554644
 KEYWORDS
 SOURCE
 ORGANISM
 Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 REFERENCE
 AUTHORS Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
 TITLE Male tissue-preferred regulatory region and method of using same
 JOURNAL Patent: WO 0160997-A 9 23-AUG-2001;
 PIONEER HI-BRED INTERNATIONAL, INC. (US)
 FEATURES
 source
 1..255
 /organism="Zea mays"
 /mol_type="unassigned DNA"
 /db_xref="taxon:4577"

Query Match 17.5%; Score 243.6; DB 6; Length 255;
 Best Local Similarity 98.4%; Pred. No. 1.3e-38;
 Matches 246; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1145 CGGAGCTTCGTCATCTCATGCAATGCAATGCTAGTGTGTTCAACGGTTCGTTGT 1204
 |||||||
 DB 5 CGGAGATCCGTCATCTCATGCAATGCAATGCTAGTGTGTTCAACGGTTCGTTGT 64
 |||||||
 QY 1205 TCCATCTCCCAAGCCTTGCTATTTCTGAACCAAGAGATACTACTTCCCAACATTCAT 1264
 |||||||
 DB 65 TCCATCTCCCAAGCCTTGCTATTTCTGAACCAAGAGATACTACTTCCCAACATTCAT 124
 |||||||
 QY 1265 CTTACTCATGCAACTTCCATGCAACGACATATGTTTCTTGAACAGATCTATTAAG 1324
 |||||||
 DB 125 CTTACTCATGCAACTTCCATGCAACGACATATGTTTCTTGAACAGATCTATTAAG 184
 |||||||
 QY 1325 ATCACAACAGCTAGGCTTCGCCGTAAGCTTCCCTCTCTGCTGCGATCTTTTTCGTC 1384
 |||||||
 DB 185 ATCACAACAGCTAGGCTTCGCCGTAAGCTTCCCTCTCTGCTGCGATCTTTTTCGTC 244
 |||||||
 QY 1385 CACCAACCATG 1394
 |||||||
 DB 245 CACCAACCATG 254
 |||||||

RESULT 7
 AX224396 158 bp DNA 1linear PAT 10-SEP-2001
 LOCUS AX224396
 DEFINITION Sequence 3 from Patent WO0160997.
 ACCESSION AX224396
 VERSION AX224396.1 GI:1554638
 KEYWORDS
 SOURCE
 ORGANISM
 Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 REFERENCE
 AUTHORS Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
 TITLE Male tissue-preferred regulatory region and method of using same
 JOURNAL Patent: WO 0160997-A 3 23-AUG-2001;
 PIONEER HI-BRED INTERNATIONAL, INC. (US)
 FEATURES
 source
 1..158
 /organism="Zea mays"
 /mol_type="unassigned DNA"
 /db_xref="taxon:4577"

ORIGIN
 Query Match 10.5%; Score 146; DB 6; Length 158;
 Best Local Similarity 99.4%; Pred. No. 4.9e-19;
 Matches 157; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 1155 CGTGTCACTCATGCAATGCAATGCAATGCTGTTGTTCAACGGTTCGTC-TTGTTCATGTC 1213
 |||||||
 DB 1 CGTGTCACTCATGCAATGCAATGCAATGCTGTTGTTCAACGGTTCGTTGTTGTTCAATGTC 60
 |||||||
 QY 1214 CAAGCTTGCTCATTTCTGAACCAAGAGATACTACTTCCCAACATCATTTACTCAT 1273
 |||||||
 DB 61 CAAGCTTGCTCATTTCTGAACCAAGAGATACTACTTCCCAACATCATTTACTCAT 120
 |||||||
 QY 1274 GCACTTCCATGCAACGACGACATATGTTTCTGTAAC 1311
 |||||||
 DB 121 GCACTTCCATGCAACGACGACATATGTTTCTGTAAC 158
 |||||||

RESULT 8
 AC147602 186199 bp DNA 1linear HTG 17-AUG-2004
 LOCUS AC147602
 DEFINITION Zea mays clone ZMBRC0334A01, *** SEQUENCING IN PROGRESS ***
 ACCESSION AC147602
 VERSION AC147602.5 GI:51315585

KEYWORDS
SOURCE
ORGANISM

HTG; HTGS_PHASE2; HTGS_FULFROP; HTGS_ACTIVEFIN.

Zea mays

REFERENCE

Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoidae; Andropogoneae; Zea.

AUTHORS

1 (bases 1 to 186199)
Bjren, B., Nusbaum, C., Lander, E., Butler, E., Wing, R., Bharti, A.K. and Messing, J.

TITLE

Zea mays, clone ZMMB0334A01
Unpublished

JOURNAL

REFERENCE

2 (bases 1 to 186199)
Bjren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,

AUTHORS

TITLE
JOURNAL
REFERENCE
AUTHORS

Anderson, M., Azechi, H.M., Barne, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Fero, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kama, A., Karatas, A., Keller, C., Lander, E., Levine, R., Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachpka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Teffaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (19-DEC-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 186199)

Bjren, B., Nusbaum, C., Lander, E., Butler, E., Wing, R., Bharti, A.K., Messing, J., Abouelleil, A., Allen, N., Anderson, M., Anderson, S., Archach, H.M., Barne, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Fero, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kama, A., Karatas, A., Keller, C., Lander, E., Levine, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachpka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Teffaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Submitted (17-AUG-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Aug 17, 2004 this sequence version replaced gi:4965659.

All repeats were identified using RepeatMasker:

http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@road.mit.edu

Bharti, A.K. and Messing, J.: The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers, The State University of New Jersey, 190 Frelinghuysen Road, Piscataway, NJ 08854, USA

(http://gsir.rutgers.edu)

Butler, E. and Wing, R.: Arizona Genomics Institute, Biological Sciences West, 448A, P.O. Box 210088, University of Arizona, Tucson, AZ 85721, USA (http://www.genome.arizona.edu)

Project Information

Center project name: L30003

Center clone name: 334 A.1

Consensus Information

This consensus is derived from a shotgun assembly that has been manually curated. It is the best representation of the BAC that we can generate without further laboratory work. The draft assembly has been edited, and if possible, ends identified by vector as well as by BAC end sequences, and contigs ordered and oriented. Bases that are not Ns are either above Q20 or manually edited. This assembly was performed with Arachne (Genome Res. 2002 12: 177-189; Genome Res. 2003 13: 91-96). All trace files for this project are available at the NCBI trace repository (http://www.ncbi.nlm.nih.gov/Traces/trace.fcgi?). An exact list of reads used in this assembly are available at http://www.broad.mit.edu/annotation/plants/maize/randomclones.html.

* NOTE: This is a 'working draft' sequence. It currently consists of 6 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.

1 100617: contig of 100617 bp in length
* 100618 100717: gap of unknown length
* 100718 104730: contig of 4013 bp in length
* 104731 104830: gap of unknown length
* 104831 115104: contig of 10274 bp in length
* 115105 115204: gap of unknown length
* 115205 156396: contig of 41192 bp in length
* 156397 156496: gap of unknown length
* 156497 179936: contig of 23440 bp in length
* 179937 180036: gap of unknown length
* 180037 186199: contig of 6163 bp in length.

FEATURES

source
1..186199
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone_lib="ZMMB0334A01"
/clone_lib="CHORI-201 Maize B73 BAC MboI Library"

ORIGIN

Query Match 5.5%; Score 77; DB 2; Length 186199;
Best local similarity 71.4%; Pred. No. 2.5e-05;

Matches 115; Conservative 0; Mismatches 45; Indels 1; Gaps 1;

DB 780 CTTTCTGATTTTAAAGCTAGTTGGCAACCTGTTCTTCAAGAAATTTGATT 839
|||||
DB 162058 CTTTCTTCTTACTTTTAAAGGCTAGTTGGCAACCTGATT-TTCTAAGAAATTTTCATT 162116
|||||
QY 840 TTTTCAAAAAATTTGTTATTTTCTCTTAAATTAACAACTTGAAAAATAGAG 899
|||||
DB 162117 TTTCAAGTAAATTTGTTATTTTCTTCAACAAATTAACAAATTTATTTGTTAAATAGAG 162176
|||||
QY 900 TTGCGAGTACCTCCTAGATGTTTCCCAATTAATTAACA 940
|||||
DB 162177 TTTCCAACTAACCTCCTAGATGTTTCCCAATTAATTAACA 162217
|||||

RESULT 9

AC117267 25769 bp DNA linear INV 21-FEB-2004
LOCUS AC117267
DEFINITION Dicotyledonum discoidium chromosome 2 map 5836555-5862024 strain AX4, complete sequence.

ACCESSION AC117267 AC115597
VERSION AC117267.2 GI:42733680
KEYWORDS HTG
SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
REFERENCE Bukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
AUTHORS 1 (bases 1 to 25769)
Gloeckner,G., Eichinger,L., Szafarski,K., Pachbat,J., Dear,P.,
Lehman,R., Baumgart,C., Parra,G., April,J.F., Guigo R., Kumpf,K.,
Tungel,B., Cox,E., Quail,M.A., Pletzner,M., Rosenthal,A. and
Noegel,A.A.
TITLE Sequence and analysis of chromosome 2 of Dictyostelium discoideum
JOURNAL Nature 418 (6893), 79-85 (2002)
MEDLINE 22092622
PUBMED 12097910
REMARK The Dictyostelium Genome Sequencing Consortium
REFERENCE 2 (bases 1 to 25769)
AUTHORS Baumgart,C.
TITLE Direct Submission
JOURNAL Submitted (09-APR-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
3 (bases 1 to 25769)
REFERENCE Baumgart,C.
TITLE Direct Submission
JOURNAL Submitted (21-FEB-2004) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
On or before Feb 21, 2004 this sequence version replaced
gi:19570016, gi:20087114.
CDS predictions from GeneID may contain errors. Further Information
is available from IMB Jena, Department of Genome Analysis
(http://genome.imb-jena.de/dictyostelium/)
and the University Cologne, Institute for Biochemistry I
(http://www.uni-koeln.de/dictyostelium/project.shtml
Funding
Agency : Deutsche Forschungsgemeinschaft (DFG).
Location/Qualifiers
1. 25769
/organism="Dictyostelium discoideum"
/mol_type="genomic DNA"
/strain="AX4"
/db_xref="taxon:44689"
/chromosome="2"
/map="5836255-5862024"
complement(join(832..1070,1237..1349,1533..1798,
1896..2053,2231..2360))
/note="Geneid exon scores (in order of location ranges) :
23.68, 1.54, 15.75, 9.21, -4.75 - GSCJ_ID dd_03387"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAS38625.1"
/db_xref="GI:42733681"
/translation="MRLOIILFLSTILIRKLTSTSSNDISTLITPFPKYNNO
ILSKSEMFRLSESSNKYIFEFDFGNHREYBNDRIKERENKVGVDSPFCL
NSKPNLEVIKKEFTISRNNKVPENIWMISTHENCYNSNYLYNNKNNYSQELV
FKQVAKSFLQNGYNIGIKSSKNTPFKIFENKDKLFSLPFYKNTIAENDEFDYY
SSPLSGMTDFKIKQFKKLDIETILSDVEVNLCSNNNNNNNNNNVDEDTLLGSGSG
SSGSGSGSGTSGAOLSGTGYVNHQ"
join(2957..3173,3272..5583)
/note="Geneid exon scores (in order of location ranges) :
24.88, 254.96 - GSCJ_ID dd_00753"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAS38626.1"
/db_xref="GI:42733682"
/translation="MEFPKYVYITNSLDKPKEMWTQSMSTLTITTDITQKFPVNSG
SSSSSKGFEAVLDLDTAGQEEYSAMDQVKTGDCPMIVFSIDRSSEFEVSQLK
HEIRVKRQDDVPIIVGNKVDLESRRQVRIEADQLASLRVPIETSAKTRRSIEE
AFETLRVTPRNTVYKVVMGGGKSAIIIOFIQNHVEVEYDPIEDSYKQVYIS
GLPIGSLNKSSSSSSSSSSSGKTGLFNKIPSGDKQSPQQAASPTIDRIGCI
STNRLAANYLSKSMNLKREVPITGDCYCOGCAVILSRPSNLITGDDSTDMCEP
CKYSNNILLEGELPNKDSVERIVLSSPSTSTDSKREESIIICYCIDVSGSWITTT
EVPSLQSEWNAKGVKAGASGPYSYISRLCEVQSSIPMTIDRLSIQYPMKRVLLVTFPS
DEWMTYTQSNVDPGPIVAGDLIEDFDQLEIGRSWTYDKLPTAGSSSPFLAKIKLSL

EPVOSTALGPALLVSAIASOKMSEVVICTDGVNPGAGIEDLLPGAPQFEYEVKT
KLAONKTTINIGISGSHIDGVGKXSEONNGITIIHLELRREIRKLTQNPMTA
TDVENSICLHPTEINRKDSKQSRVYKQFPNPNVSLDTDILYSRRNPTEFVQIY
FOIQTIKTLDGVRCLRVASQOLQTPDNPNTSIVANISITLMAKTQAAKQOOEY
MESRLHLAAATKIRLSLCTDQEWEEFNFLEBEMAPLITCIKNQOQVEKATD
DEIQVFMKQVHKHSFVEGGRKQDISRRKGEAEINKQYNIKFT"
complement(5711..6796)
/note="Geneid exon scores (in order of location ranges) :
68.61 - GSCJ_ID dd_00752"
/codon_start=1
/product="similar to Plaeomedium falciparum (isolate 3D7).
Hypothetical protein"
/protein_id="AAS38627.1"
/db_xref="GI:42733683"
/translation="WSSRVKSLIRVYNGINLTIEKSLSPRLINNTNNNNNNNN
GTTTTTGCGNNNNNSTPSTNPVSFDFVDHSSNNQOYIVDSFVNTNNNNNNNN
NN
KFNNNNNYFPSTIEMITTLTNSCTLEDIGTIDNCYGTNTTSPSTSSSTSLAN
KSNHMMPSLKINOIINDNNNNININFEVQIOSNSSLSSCSTYOSCPIDEHIOE
OLRIQEKQKOIDDODDDDDDDDDDDDDSPPPPTTTTKTFCLSPPRIKKEENID
QNSMPPPKLDRRSRLIPTTANT"
join(8924..9076,9236..10573)
/note="Geneid exon scores (in order of location ranges) :
10.37, 128.60 - GSCJ_ID dd_00729"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAS38628.1"
/db_xref="GI:42733684"
/translation="WSSSTITTPATTTTNNNSSTNGIHNKTNLTFLKKNIKNI
PNNAIPKKNVGIIGGNVNNNNNDKRLNSDNGVFNKQILFPSPKSNMSTYKHS
SSSSSLAKNEYDIESIKELKSMIDITGEGISPLSPRESPISTIDESTATYTK
LTDAPVAASSSSSVKTYIDVLSFKSANTRPDIIVENHOKRIVPLSLPTPN
NNNNNNNNNSQNTANGOKPIEPSQIISBPKIAVQASSTKTNNNNNNNNITP
TKKANSRTTPOSNRNNHNSTKSTKQNPPIPSLNTIPNOKNTITPSKSKTA
TTPSKNKIINDVNAAPAAVANSSTTTEQVVAIIDATPSTNESPRTPLKEPDLRA
ARORIDIGPKTAGYKNTIALVPKSRKRPKTPKNQOVCSKSMDOIKKMRQLH
ENYDWTGIPNEKEVEASNNESTESTLEKEMNKLEISEENHNH"
complement(10972..11979)
/note="Geneid exon scores (in order of location ranges) :
23.29 - GSCJ_ID dd_00728"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAS38629.1"
/db_xref="GI:42733685"
/translation="MEKIKYFPIIDRONKPKIOEFTSKLIGKESPTTEKPIY
EKLAMTOSIESILKNTGINTKNSIFBEIYCYGTFLEBYRPIYVLTKTTTBEIIL
DDYLPETKNSLNMKNEKEKILNSILTYNNNEBKIGEFWEIINEFONFTKESFOR
IINNSLMIKSIDSRNINSNPINSRCSFPEKRSDDASGDFLTLSMIGIMONYI
ENSILSEFOTINYHAKSFPLINDLVSFNRBIENDLLVYIKLAIOLMSIOLSID
KTIELIDHYLKLFSIETILKQNDSTVQLKQVPQNSKILSGIYPAHKSXRY
N"
join(12884..13020,13172..15320)
/note="Geneid exon scores (in order of location ranges) :
2.42, 136.56 - GSCJ_ID dd_00727"
/codon_start=1
/product="similar to Dictyostelium discoideum (slime
mold). Hypothetical 97.7 kDa protein"
/protein_id="AAS38630.1"
/db_xref="GI:42733686"
/translation="MDNLIYDPLIKREYNQOSFLDCLTILSINGLYISYDGLDYCNN
FLIENKTVKKEALISYGLANTQYQLDEFRLLKLKKRTDCKKIIINTHSLAKCC
GNKEDVAILIIEINNNNIINKDOFNNNNNNNNNIINNNDNNNSIIEEDLAKIKI
TLKNERJINGDGEITTTTTTTSRSSTTITTTTNGVANNYLAPOWRNIVIK
TOILEHLALNINAHANKYFRLPIQLVYKFGYIQASWLVYHDDDDADSDYSIDI
DDDSYSDGNCSSSSSSSDISSSNSINNGISSSSSILSNSLSLEISMRNLSL
PEGKSIJEFEXEYVITQDRLPISGISIRFPGNORLAKVVISDNVVISITFGSFNQ
SLDGWMLPKQLKQFGHKFOQITKMGQLPESLTSLLIDPSKYGKVIIGISIPDSKT
LDVKNSSQNOEHSISPNIPISIRLVDSDFNONIKANDISSNVTLSIHGEHFN
DIGIKLPSNLSIKRGAIPORDIKLQGSSTSIDPCKGRNRPASMTQTLTSDIGS
KPNQITPGCITHTKLKSLANGYHFNQITPADTLPPTLESNLGSGREIYVNDKDEYD
CYGSLNKGKFSNSSNFCVGTGNGNLEMLKNTSLTKTLTFLANFNRIEVDGLPNSI
ESLNIQYHFNQPIGNWVLPKLLKTLFLNSEFONISADGCI PFGLOTIYIRNSMNF
INSNPLITKYINIIDLSHL"

CDS
join(16184..16199,17272..17613,17707..19331)
/note="Geneid exon scores (in order of location ranges):
0.08, 32.32, 227.42 - GSCU_ID dd_00726"
/product="similar to Dictyostelium discoideum (Slime
mold). Spore coat protein Sp87 precursor (PL3 protein)"
/protein_id="AAS38631.1"
/db_xref="GI:42733687"
/translation="MFLGEDCNKITNEECHKSECVIVNYTPCCGEQKACSGTFD
TCTYNSCSNNSNOVVEVSNKCFNLDGFIKITTPEVSCDAKIKCALCKSCSF
OKNSCSNPTSCGSESICEGLSGSSTSGSGSGSGSGSGSGSGSGSGSGS
SSGSSGCVSSTGTTGPEGHCSMNVDVATCLASTGTGTLGTSSTAGVSCSLT
LCPIGHI CVENSGVNCVPGSGTSGSGSGSGSGSGSGSGSGSGSGSGSGSG
TACVPSPTGSSSTSGGHPDCKDVTCPGPHCECKDGTACVPSPTGSSSTSGG
NTPCSNVNCPDGFYCECKDGKTAACVPSPTGPPKPVCSLCPNHECRVNDQHQ
CVKHHDRCSLRCPHGHECVNDHNGKCCVSHRPPPEVCSLCPKHECKFDDHG
KCCCVKIHCDVEVDLDGSGPECKIRHDSGRCVSRPHPPOHEKKNKRCPPGHECK
VDQHGKCCVVAHRPPKCSLRCPNHECRVNHGEECVKVAHDKCSLRCPGHECK
VDQHGKCCVVAHRPPKCSLRCPNHECRVNHGEECVKVAHDKCSLRCPGHECK
LHCAMKTPIDKXNCCERKPVVCTSNPLDGGHGF1"
19762..21336
/note="Geneid exon scores (in order of location ranges):
106.68 - GSCU_ID dd_00725"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAS38632.1"
/db_xref="GI:42733688"
/translation="MDNNNLPPQYFPPOSNNYCCDCDNDNNDSTLTIOKPIAPSPS
SOOPSTIIQIHNHKKYLFKEFEKHNNIDSMKLCQSEFISPKATNTDSMLN
SLMLPTGFKLVNDHLKKVYIPFHYGVAVVYVFGFGLDKNLNLVDYGVNTN
STYNNNNYNNNNNSLNNNNADKKTTTSTATTILNENVELNNGNNNNNNKGS
EDLKKKSIKNNVVDGTTITRTIRILILACNSDLSDESINCMWSQODENNVLEK
SLDFEQFQALDQVEVLAARYEPVWVRDNLNLTLEDSMSLEYKQOQSPFNVNV
SSTIQNEKLFSLVYIPFYTKYEFQNSQFPLNONGVISAIREIGIGKIDGLK
YTGNYSTLWNGEPEPKCKGDELAIIDSKVNTSSYFLTFRSSTNFGSISHL
TIRNNSKSIATLRGKRRSVQKGNPSSLPLHERFSPFGHMCFFLIIDENVQIDVI
KINREGNITVQISKN"
KINREGNITVQISKN
complement(21961..24357)
/note="Geneid exon scores (in order of location ranges):
173.48 - GSCU_ID dd_00724"
/codon_start=1
/product="similar to Arabidopsis thaliana (Mouse-ear
cress). Putative PRG1-like negative regulator"
/protein_id="AAS38633.1"
/db_xref="GI:42733689"
/translation="MSKIVSKSQTVSPGPKQNIPTALPSVAPSLSLQKPNNGAG
SHRKORNSGSSKNRPNLPADQIKSTERKSKSNKNNKNNKXYSKNNVSNVNN
SSSIVIKCEQO
SPSSSSSTSPSCSLSSSSSTSPQSEIAEKNDIINNKNNTNGNFEEBEERKRV
EQQSIEETPESFEETKSVVEVEPEEQFELIPISIGETLKETYSFLSVLSELIV
DMYNDDDKSSSTSSSSSSSSSSSPSPKPKIKNNNDNNNTTTTATTTTTSKTATSS

Query Match 5.3%; Score 73.4; DB 3; Length 25769;
Best Local Similarity 49.4%; Pred. No. 0.00014;
Matches 217; Conservative 0; Mismatches 221; Indels 1; Gaps 1;

CDS
438 CGAACAAGATTAAAAAATAACAAGTAACTAATCCTCGAAAGCTATCATGTAA 497
2392 CGAAACAACAGCTAAAAAATTTGTTGAGAAAAAATAAAAAAATAAAAAA 2451
498 TGTTTAAGAAACATCATTAACAACGATCTCTTAATAAAAAACAAGCATTTCCGAAG 557
2452 TATGAAAAAATAAAAAAATAAAAAAAGTTTTTACAAAAAATAAAAAATATCTAAAAA 2511
558 AGACAATATGTTACAGTTACAAACATCTAAGAGCGAACAATTTATGAAAGGTAA 617
2512 AATTATTTTATTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTA 2571
618 CTATGACGTTCAAGATTTTCTTTTCACTCTGTATTTTGTATTTGTTTATATACAT 677
2572 ATAAAAATATTAATTTAATTTAATTTAATTTTCTATTTTATTTTATTTTGAATAAT 2630
678 TTTCTTCTTACAAATAGAGTATTTTCTTCGATTTTATTAATAAGTAAAGTCAAT 737

Db 2631 TAAATTTTAAAGATATTTGTTTTTTTTTTTTTTTTTTTTTTTTTAAAGTCAATTT 2690
Qy 738 TTTATTAAGAGACCAAGCATGCTAGATTTCTGTTCAAAAATCTTCTGATTTTTTAAG 797
Db 2691 TTTAAATATGGGTTGTTGATTTGAAAAATTTCTTTAAATTTTTCAGGAATATCTTTT 2750
Qy 798 AGCTAGTTGGCAACCTGTTCTTTCAAGAATTTGATTTTTCAAAAAATAATAGTT 857
Db 2751 TTTTTTTTCAACAACCAACCAATTTTTTTTTTTTTTTTTTTTTTAAAAAACCATAATT 2810
Qy 858 TATTTCTCTTATTAATAT 876
Db 2811 TTTATTTTATTTATTTTAT 2829

RESULT 10
AF034389/c 713 bp DNA linear INV 04-FEB-1999
LOCUS
DEFINITION
Plasmodium falciparum sexual stage antigen (816) gene, promoter and
partial cde.
ACCESSION
AF034389
VERSION
AF034389.1 GI:3098290
KEYWORDS
ORGANISM
Plasmodium falciparum (malaria parasite P. falciparum)
Bukaryoca; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE
1 (bases 1 to 713)
König, R.N. and Stunnenberg, H.G.
Isolation and functional characterization of two distinct
sexual-stage-specific promoters of the human malaria parasite
Plasmodium falciparum
Mol. Cell. Biol. 19 (2), 967-978 (1999)
JOURNAL
MEDLINE
99108072
PubMed
9891033
REFERENCE
2 (bases 1 to 713)
Decherter, R.J., Kaan, A.M. and Königs, R.N.H.
Direct Subcloning
Submitted (12-NOV-1997) Molecular Biology, University of Nijmegen,
Toernooiveld 1, Nijmegen 6525 ED, The Netherlands
FEATURES
source
1..713
/organism="Plasmodium falciparum"
/mol_type="genomic DNA"
/strain="NF54"
/db_xref="taxon:5833"
1..>713
/gene="816"
1..536
/gene="816"
537..>713
/gene="816"
537..710
/gene="816"
711..>713
/gene="816"
/gene="816"
/codon_start=1
/product="sexual stage antigen"
/protein_id="AF034389.1"
/db_xref="GI:3098291"
/translation="M"

ORIGIN
Query Match 5.1%; Score 71.2; DB 3; Length 713;
Best Local Similarity 45.8%; Pred. No. 0.00045;
Matches 282; Conservative 0; Mismatches 333; Indels 1; Gaps 1;

Qy 444 CAAGATTTTAAAAAATAACAAGTAACTATCACTCGAAGCTATCATGTATGTTTA 503
Db 681 CAAAAGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 622

Db 612 TTGTAATTTTGTATTTTAAATATTTATATACATAAAATACCTTTAAACACAC 553
Qy 712 TTTTATTAATGACTATAAGTCACTTTTATATAGACGACGATGCGAGATTCGCT 771
Db 552 ACACACACACACCAATGTTTTTTTTTGTGTTTTTTTTGTTTTTTTATCAACCT 493
Qy 772 TCAAAATCTTCTGATTTTTTATAGAGACTAGTTGGCAACCCGTCTTTCAAGAT 891
Db 492 TTCATATGCTTTGATTTAGTTATTTTTTTTTTTTTTTGTAATGTTTTTTTGA 433
Qy 832 TTTGATTTTTTCAAAAAATAGTTATTTCTCTTAT 871
Db 432 ATTTTATTTTATTTATTTATTTTATTTTATTTATTTT 393

RESULT 12
LOCUS BX957346/c 149526 bp DNA linear HTG 08-OCT-2004
DEFINITION Danio rerio clone CH211-117K16, WORKING DRAFT SEQUENCE.
ACCESSION BX957346
VERSION BX957346.13 GI:54019944
KEYWORDS HTG; HTGS_PHASE2; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Danio rerio (zebrafish)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 149526)
Beasley, H.
Direct Submission
Submitted (07-OCT-2004) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfish-help@sanger.ac.uk
zfish-help@sanger.ac.uk clone requests: clonerequests@sanger.ac.uk
On Oct 9, 2004 this sequence version replaced gi:53850295.

COMMENT
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: zc117K16
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 149526 bases at least Q40
Consensus quality: 149526 bases at least Q30
Consensus quality: 149526 bases at least Q20
Insert size: 149526; sum-of-contigs
Insert size: 149176; 6.5% error; agarose-fp
Quality coverage: 12.41x in Q20 bases; sum-of-contigs Quality coverage: 12.54x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES
source
1. 149526
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="CH211-117K16"
/clone_id="CHOR1-211"
1. 149526
/note="assembly_fragment:03115
clone_end:SP6
vector_side:right
clone_end:T7
vector_side:left"

ORIGIN
Query Match 5.1%; Score 70.8; DB 2; Length 149526;
Best Local Similarity 46.8%; Pred. No. 0.00044;

Matches 260; Conservative 0; Mismatches 292; Indels 4; Gaps 1;
Qy 488 TATCATGTATGTTTAAAGAAACATCTTATTAACACGATCCTTTAAAAACAACAT 547
Db 109063 TTTTAAATATATGAAACAAAATCATATTTTACTTTGCGCTTTGATGATTA 109004
Qy 548 ATTTCAGAAAGACAAATATATGTTTACA-----GTTTCAAAACATCTAAGAGCGCAATTA 603
Db 109003 GTTTTAAAGATTTTAAAAATATTTTAAACATGTTTATATACATATATATGACATTAAGAG 108944
Qy 604 TATCGAAGGTAGCTATGACGTTGAGATTTTCTTTTCATTTCTGTTATTTTGTAT 663
Db 108943 CCAAAAGTACTATATGCAATTTTAAATGTTTTTATTTTGAATTTTTTTTAT 108884
Qy 664 GTTTTATATACATTTTCTTCTTCAATATGAGATATTTCTCCGATTTTATTAAT 723
Db 108883 AT 108824
Qy 724 ACTATTAAGTCAATTTTATATAGACGACGACATGCTGATTTCTGTTCAAAATCTTT 783
Db 108823 TATATATTTATTAATTTATATATATATATATATATATATATATATATAT 108764
Qy 784 CTGATTTTAAAGAGCTAGTTGGCAACCTGTTCTTTCAAGATTTGATTTTTC 843
Db 108763 ATATATCATATTAATAAATCTTTATATATGTTATATATATATATATATATAT 108704
Qy 844 AAAAAAATTAAGTTATTTTCTTTTATATATATATATATATATATATATATATAT 903
Db 108703 AGAAAAACAAAAGTTATATATATATATATATATATATATATATATATATAT 108644
Qy 904 CAGACATGAGCTAGATGTTTCCCAATTAATTAACAATGATATATATATTTGGC 963
Db 108643 TATATATGCGCTTCAATGATCTTATATTTTAAATATATATATATATATAT 108584
Qy 964 AGCCCATTAATTTATTAACCGAACTGAAATCGACCAACCAATCTGAGCTATTC 1023
Db 108583 ACATATATATATTAAGAAAGACCAAAAGTAAAAAATGTTATTTATATCTGA 108524
Qy 1024 TCTAGATTATGTAATA 1039
Db 108523 GTTTTATATGTTAAAA 108508

RESULT 13
PFMAL4P1/c 347582 bp DNA linear INV 29-JAN-2003
LOCUS Plasmodium falciparum MAL4P1.
DEFINITION AL034557 AL844503
ACCESSION AL034557.8 GI:23498126
VERSION
KEYWORDS
SOURCE Plasmodium falciparum 3D7
ORGANISM Plasmodium falciparum 3D7
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
1
REFERENCE
AUTHORS Hall, N., Pain, A., Berriman, M., Churcher, C., Harris, B., Harris, D., Mungall, K., Bowman, S., Atkin, R., Baker, S., Barron, A., Brooks, K., Buckee, C.O., Burrows, C., Cherevach, I., Chillingworth, C., Chillingworth, T., Christodoulou, Z., Clark, L., Clark, R., Corton, C., Cronin, A., Davies, R., Davis, P., Dear, P., Dearden, F., Doggett, J., Felwell, T., Goble, A., Goodhead, I., Gilliam, R., Hamlin, N., Hance, Z., Harper, D., Hauser, H., Hornby, T., Holroyd, S., Horrocks, P., Humphray, S., Jagels, K., James, K.D., Johnson, D., Kerhornou, A., Knights, A., Kontorov, B., Kyes, S., Larke, N., Lawson, D., Leonard, N., Line, A., Maddison, M., McLean, J., Mooney, P., Moul, S., Murphy, L., Oliver, K., Ormond, D., Price, C., Quail, M.A., Rabbinowitch, E., Rajandream, M.A., Rutter, S., Rutherford, K.M., Sanders, M., Simmonds, M., Seeger, K., Sharp, S., Smith, R., Squares, R., Squares, S., Stevens, K., Taylor, K., Tivey, A., Unwin, L., Whitehead, S., Woodward, J., Sulston, J.E., Craig, A., Newbold, C. and Barrett, B.G.
TITLE Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13
JOURNAL Nature 419 (6906), 527-531 (2002)
MEDLINE 2255708

[illegible]

KKKTKSTIDILIRVINIEKSDYDITFKLSPNRRIYPISGKYRKGRRIYILEBDSGTDSCG
 IDHVSISSESESESEEMEDINDIYAPRAPAKITILEVVLSPSGNNITASANNITASGN
 NITLASGNNITASGNTPESDIQNDIQSDGISPSKTKTLENNITLKDPEFSQYIOSEPDP
 VPNDSSDIDPEPNTQHTLYPDKDEKPEFYSIDRLLYGEENYXNMSNSGNDLYX
 NGKNNLYSGONNYSIGDITPSYDNRGLKSGHDSYSGDILNDPLSGNHQIDIDYEVLX
 KKKENLPEOTNHVAKTITTRNRPKAPRBDPLHQLELPEFTYILDRHRNCEKKNKEELD
 KLKWEEMETHSGNTHPSDSKRTLNTVYSIQIDMDNRPINQPFNMDINDVPTMDNM
 EDDIYDVNDNDSDTSTVSNMDVPSKVQIEMDVNTKLVKERYPIADWDI"

45458. .46601
 /gene="PF00010w"
 /note="VAR exon II, could be alternate exon II for upstream VAR gene PF00005w"
 complement {join(48479. .49537, 49768. .49821)}
 /gene="RIF"
 /note="RIF"
 /gene="PF00010w"
 /note="VAR exon II, could be alternate exon II for upstream VAR gene PF00005w"
 complement {join(48479. .49537, 49768. .49821)}
 /gene="RIF"
 /note="synonym: PF00015c"
 complement {join(48479. .49537, 49768. .49821)}
 /gene="RIF"
 /note="1 probable transmembrane helix predicted by TMHMM2.0 at aa 328-350
 Similar to Plasmodium falciparum rifin pf00030C
 SMALL:096112 (EMBL:AE001367) (370 aa) fasta scores: E():
 3.4e-65, 53.6% id in 375 aa"
 /codon_start=1
 /product="RIFin"
 /protein_id="CAD49095.1"
 /db_xref="GI:23498128"
 /db_xref="UniProt/TREMBL:Q81219"
 /translation="MKIHYINILIFELPLNITLYNORNHKSTNLPHONNSLCEBELY
 APAYDDPQMKVEYVWFSKQTOQRFHYEYERWERMKQKDCDEKIQKILIKDNE
 KALMDKPEITQTDIOSDAIPTCVEKSLADKVEVSGIGATVAPYGLIGSVAV
 HWMPKALAEALIAKALAEGLADILAAALAEAGKRGMEFVILKALIKFVEVAPGIGI
 ISNGYNTKAEFVNTYISKRYNGTGNRBSINPTLAETIELKISFGVNGGDDHP
 LYAIRQMKGLAEATPAKAAAPAKAKATPAIKKEQOTLIERGNSITSINASTII
 AILILVITVITVILYILYRRKKKKKKLOYLKLE"
 complement {join(52002. .53297, 54200. .63307)}
 /gene="VAR"
 /note="synonym: PF00020c"
 complement {join(52002. .53297, 54200. .63307)}
 /gene="VAR"
 /note="Similar to Plasmodium falciparum variant-specific
 surface protein var-3 SMALL:Q26032 (EMBL:LA0609) (3006 aa)
 fasta scores: E(): 6.2e-107, 43.31% id in 3551 aa"
 /codon_start=1
 /product="erythrocyte membrane protein 1 (PEEMP1)"
 /protein_id="CAD49096.1"
 /db_xref="GI:23498129"
 /db_xref="GOA:Q81218"
 /db_xref="UniProt/TREMBL:Q81218"
 /translation="MGTGSSPSVPRKQVKNKSHNSARVLENIGETIYNEKKKNQY
 TSOALRGDLSRARFHDGLRKARLGVITPGPANSCLDHLKPYTNINNGYPRAPNCDLNR
 ONLVATGENEAICNSDKLRVTKGSAGGCAPIFRQNRCDNLELMDITNDITDDDLG
 NNRVATYRGGSIYAKHPKENSESVCTALASPADIGDIVRGKMEFRNEDVAVQKL
 RAVPKETINDILKEKEISDYDNDPNYVYILREDWWTQWRAITCYIPIYVYVFKT
 SDDTIVFNIDKCGHYGAPPTNLDYVQPLRPMDEAAEFECIRHMKAKIICDAGCN
 DKBLTQSNQYDCTKRIKGGSCSRENTCGSCNVDVPMLEKQONEFQIOXKY
 DKSEIETVYNTKPISTNSNNTKKEYEYFEYELKKQYSVQNPOLANNGYCEKLEB
 EDALDIFTKTKKHAIFRSDYQCPDVCVDCDEGTCYQKTDIDDCDSKTIYQILISE
 TPRIEVLVYSDQKQVITKGLKDFCRPNYNNENLQMKYKNSYKNSYKNCENISMLY
 QDPKPEYNSIVSECFHSMARKLIDITIRRHQULKCNNTNNTVDCISKECNCEYA
 MIEKKDQMEWELKVELNKDSEHSYNNYKQKLDVDRLLFQWMLALDDDEKGDQFPE
 DLKKKPGVSEVAGTANSODALIEFLDLKMDNLTDQDNNISITPCYPRPPTPGCT
 NNGGKLYRVRGLAEMQBARROLERGGSEINLKDASGAKTIRGKEKKLNGQICN
 IDTYSISDSNRNGNPGCTGKNDGRFETIYEWYSIGHEKRRTHPEYVMPRRRLHCIS
 NIELKDVSVYIKGNASHSLGDVLLAAKTEAKNIKELYOONNSKGVINDQNKETIC
 RAMKYSFWDIEDIDIRGDMWQNTDATKLOAYLAKIPDKIKDNHKKIKGLQYNGTD
 HKLIRBDMWENHRQVRAWKCALIENDKMKNGIPIEDYIPORLMMWMAWYCKE
 OSRLYGBELKSCQCKRQCTBGSDVDCGCKYACDQRYDNEINKREBPYTKINGKTT
 LYKATKRGVTSNNPDKQVVDVFLKQYLPRKSKTPGVATNPVTLVSSAGAYVHQ
 ELGVTQKQVTSNNPDKQVVDVFLKQYLPRKSKTPGVATNPVTLVSSAGAYVHQ
 LLDNSNGTQGLDGCNPKIGNYSNMENRSELAENKAGVAPPRREKFCVSLAKXGI
 FKNQGEIDREFVYSALLETYFAMKRNNDQNKAAEELKSGTIPENFKRQMYTLPADY

Db 273 NTTT 270

RESULT 15
BV119878/c 241 bp DNA linear STS 19-MAR-2004
LOCUS PA01377 CML247 Zea mays CML247 Zea mays STS genomic, sequence
DEFINITION tagged site.

ACCESSION BV119878
VERSION BV119878.1 GI:45589251
KEYWORDS STS.
SOURCE Zea mays
ORGANISM Zea mays

REFERENCE Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
TITLE Clade: Panicoideae; Andropogoneae; Zea.
JOURNAL 1 (bases 1 to 241)
McMullen,M.D., Vroch Bi.,I., Schroeder,S.S. and Gaut,B.S.
MPZ-UCI Joint SNP Discovery
COMMENT Unpublished (2003)

Contact: Brandon S. Gaut
Dept. Ecology and Evolutionary Biology
U.C. Irvine
321 Steinhaus Hall, Irvine, CA 92697-2525, USA
Tel: (949) 824-2564
Fax: (949) 824-2181
Email: bgaut@uci.edu
Primer A: TTCTAGCTCAATGATCATCAAGA
Primer B: AGGAAACCAAGCCTTATGAGAGA
STS size: 241
Protocol:
PCR amplification of genomic DNA
Template: 50 ng
Primer: each 0.5 uM
dNTPs: each 200 uM
Taq Polymerase: RedTaq (Sigma)
Total Vol: 10 uL
Amplicon sequencing
ABI protocol - using d-Rhodamine terminator cycle
sequencing ready reaction with amplitaq DNA polymerase FS
Sequence ran on ABI 3700 sequencer.

Buffer:
Genomic DNA amplification
RedTaq (Sigma)
Sequencing buffer
d-Rhodamine kit (ABI)

PHRED/PHRAP Quality Scores 56 56 56 51 51 40 40 40 40 40 35 35 42 48
56 56 56 56 56 42 42 56 33 33 33 47 44 37 37 37 37 37 48 48
42 42 42 42 30 30 30 37 37 44 44 44 84 90 89 81 71 68 51 40 45
45 51 39 39 39 56 44 48 86 86 90 81 71 68 63 62 58 58 69 69
67 67 77 72 68 71 68 72 74 80 74 76 73 66 77 79 68 62 57 56 53 60
58 71 71 68 68 77 75 79 71 79 68 81 75 90 79 35 37 33 39 42 42 35
37 30 38 20 20 48 40 44 40 33 42 37 37 74 68 60 50 54 54 51 51 55
49 50 57 57 72 82 90 90 88 88 79 74 74 67 64 62 53 50 52 59 55
65 71 71 84 84 84 88 85 90 84 84 88 88 90 80 80 79 83 83 86 86 81
76 68 49 48 39 39 42 49 33 42 29 24 24 14 14 14 32 44 40 48 48 90
84 84 84 87 79 72 72 63 57 53 51 41 57 54 56 56 49 51 30 33 33 28
28 28 33 40 42 48 37 37.

FEATURES
Source location/Qualifiers

1..241
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="CML247"
/db_xref="taxon:4577"
/clone_lib="Zea mays CML247"
/dev_stage="seedling"
/note="Organ: leaf; genomic DNA from inbred line"
<1..>241

Query Match 5.0%; Score 70; DB 11; Length 241;
Best Local Similarity 61.5%; Pred No. 0.00082;
Matches 112; Conservative 0; Mismatches 70; Indels 0; Gaps 0;
QY 799 GCTAGTTGGCAACCCCTGTTCTTCAAGATTGATTTTCAAAAAATTAGTT 858
DB 195 GTTAGTTGGAAACCAATTTTCCCAAGGAAATTTCCCAAGGAAATTAGTTTC 136
QY 859 ATTTTCTCTTTATTAATTAAGAAACATTAGAAAAATAGAGTTGCACTAGCCCTAGA 918
DB 135 ATTTTCCCTTGGAAATTAAGAAATCCCAAGGAAATAGAGTTCCCAACTAACCTTAAT 76
QY 919 ATGTTTCCCAATTAATTACATCACTGTATTAATTTTGGCCAGCCCATTAATTAT 978
DB 75 ACTGTTAATTAGTTCCATATCCATACCAAGAGTAGATATCTTAACCGCATATCTGT 16
QY 979 TT 980
DB 15 TT 14

Search completed: September 16, 2005, 03:01:32
Job time : 4680.65 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

CM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 08:10:17 ; Search time 674.892 Seconds
(without alignments)
12227.321 Million cell updates/sec

Title: US-10-713-381-2

Perfect score: 1394
Sequence: 1 ccatgtgtctctcatgaaga.....cttttcgtccaccaccatg 1394

Scoring table: IDENTITY_NUC
Gapop 10.0 ; Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980a:*
2: geneseqn1980b:*
3: geneseqn2000a:*
4: geneseqn2001a:*
5: geneseqn2001b:*
6: geneseqn2002a:*
7: geneseqn2002b:*
8: geneseqn2003a:*
9: geneseqn2003b:*
10: geneseqn2003c:*
11: geneseqn2003d:*
12: geneseqn2004a:*
13: geneseqn2004b:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1394	100.0	1394	2	AAX07409 Zea mays
2	1394	100.0	1394	5	AAH76333 Z. mays M
3	1389.2	99.7	1394	2	AAX07408 Zea mays
4	1389.2	99.7	1394	5	AAH76332 Z. mays M
5	243.6	17.5	255	5	AAH76340 Z. mays M
6	146	10.5	158	5	AAH76334 Z. mays M
7	70.2	5.0	883	4	AAH15210 Human bre
8	70.2	5.0	960	11	ACN56231 Breast ca
9	69.8	5.0	13400	13	ADR04286 Corn FT h
10	67.2	4.8	2657	2	AAZ10551 DNA seque
11	66.4	4.8	6027	2	AAZ10551 DNA seque
12	66.4	4.8	6027	8	ABX09935 DNA encod
13	66.4	4.8	6027	12	ADK12106 cDNA encod
14	66.2	4.7	439	8	ABX35844 Bovine ES
15	65.6	4.7	346	4	AAI87279 Human pol
16	65.6	4.7	8056	8	ABZ10246 Haematopo
17	64.6	4.6	9859	13	ADSS9714 Oligonuc
18	64.4	4.6	9859	13	ADSS9714 Oligonuc
19	64	4.6	8056	8	ABZ10100 Haematopo
20	63.8	4.6	431	8	ABX44556 Bovine ES

C	21	63.6	4.6	335913	5	AAI61371	Ma161371 Soybean 2
C	22	63.6	4.6	335913	5	AAI61372	Ma161372 Soybean 2
C	23	63.4	4.5	12968	4	AA545494	Ma545494 Chemical
C	24	63.4	4.5	12968	4	AA546779	Ma546779 Tumour su
C	25	63.4	4.5	12968	6	ABL34085	Ab134085 Human imm
C	26	63.4	4.5	12968	6	ABK28424	Abk28424 DNA trans
C	27	63.4	4.5	12968	6	ABN80285	Abn80285 Human che
C	28	62.4	4.5	10762	6	ABL34113	Ab134113 Human imm
C	29	62.4	4.5	10762	6	ABL32322	Ab192322 Chemical
C	30	61.4	4.4	9652	6	ABL2909	Ab132909 Human imm
C	31	60.6	4.3	5822	6	ACN56344	ACN56344 Cotton an
C	32	60.6	4.3	5822	6	ABL33097	Ab133097 Human imm
C	33	60.4	4.3	500	5	AD173089	Ad173089 Human ova
C	34	60.4	4.3	500	5	AD138222	Ad138222 Human ova
C	35	60.2	4.3	5818	4	AA546636	Ma546636 Tumour su
C	36	60.2	4.3	8896	6	ABN80326	Abn80326 Human che
C	37	59.4	4.3	461	5	ADL43834	Adl43834 Human ova
C	38	59.4	4.3	14919	4	AA546506	Ma546506 Tumour su
C	39	59.2	4.2	19459	6	ABK31213	Abk31213 Signal tr
C	40	59.2	4.2	19459	6	ABL70528	Ab170528 Chemical
C	41	59	4.2	517	13	ACN56273	ACN56273 Cotton an
C	42	58.8	4.2	5303	6	ABL32871	Ab132871 Human imm
C	43	58.8	4.2	393	8	ABX39417	Abx39417 Bovine ES
C	44	58.6	4.2	392	5	ADL43653	Adl43653 Human ova
C	45	58.6	4.2	529	13	ACN56029	ACN56029 Cotton an

ALIGNMENTS

RESULT 1	AAX07409	standard; DNA; 1394 BP.
ID	AAX07409	
XX		
AC	AAX07409;	
XX		
DT	08-JUN-1999	(first entry)
XX		
DE	Zea mays Ma45 male tissue-preferred regulatory region.	
XX		
MM	Ma45; male; tissue-preferred; regulatory region; plant cells;	
KM	plant tissue; differentiated; hybrid seed; fertility; ss.	
XX		
OS	Zea mays.	
XX		
PN	MO9859061-A1.	
XX		
PD	30-DEC-1998.	
XX		
PF	19-JUN-1998;	98MO-US012895.
XX		
PR	23-JUN-1997;	97US-00880499.
XX		
PA	(PION-) PIONEER HI-BRED INT INC.	
XX		
PI	Albertsen KC, Fox TW, Garnaat CW, Huffman GA, Kendall TL,	
XX		
DR	WPI; 1999-105628/09.	
XX		
PT	New nucleic acid encoding a Ma45 male tissue-preferred regulatory region	
PT	- useful in mediating plant fertility, especially hybrid seed production.	
XX		
PS	Claim 3; Page 23-24; 39pp; English.	
XX		
CC	The sequence is that encoding an Ma45 male tissue-preferred regulatory	
CC	region. It may be used in the construction of a vector for a method of	
CC	producing exogenous genes in a male tissue-preferred manner, which is	
CC	useful in restoring or conferring fertility, such as in hybrid seed	
CC	production. In conferring fertility, a monocot/dicot plant is transformed	
CC	with the exogenous nucleotide sequence (a male sterility gene, preferably	
CC	Ma45), which encodes a product selected from auxins, rolB and diptheria	
CC	toxin. Hybrid seeds are produced by cross-pollinating maize male fertile	
CC	and infertile plants	

XX Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other:
 Query Match 100.0%; Score 1394; DB 2; Length 1394;
 Best Local Similarity 100.0%; Pred. No. 2,2e-291;
 Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCATGGGTCTCTATGAAAAAGATGATACATGTGTCATATCCGTTTCTTAAGGTC 60
 DB 1 CCAATGGTCTCTATGAAAAAGATGATACATGTGTCATATCCGTTTCTTAAGGTC 60

QY 61 CTTCTTCGCTTATTAATGATCGAATCGGGTTACAAAAACTTCCACGGTGATAT 120
 DB 61 CTTCTTCGCTTATTAATGATCGAATCGGGTTACAAAAACTTCCACGGTGATAT 120

QY 121 CTGCAATGTCACCTTCTCCACCTCGGTGACATTTCTTGATGTGGTGGTCCAT 180
 DB 121 CTGCAATGTCACCTTCTCCACCTCGGTGACATTTCTTGATGTGGTGGTCCAT 180

QY 181 CTGACCGAGGCCCATCGACACCTTTCGGGACCCCATCAAGGGCCCTTTCGATGGCCA 240
 DB 181 CTGACCGAGGCCCATCGACACCTTTCGGGACCCCATCAAGGGCCCTTTCGATGGCCA 240

QY 241 CGAAGCTATCGGGTGTGTGTATCCAGGGATATATGTCCCCACATGTCACCTATA 300
 DB 241 CGAAGCTATCGGGTGTGTGTATCCAGGGATATATGTCCCCACATGTCACCTATA 300

QY 301 TTATATCTCTTATGATATATATTTTGGAAAAATACAAACTTATCTTTGTGTA 360
 DB 301 TTATATCTCTTATGATATATATTTTGGAAAAATACAAACTTATCTTTGTGTA 360

QY 361 GGGCTCAGCATAGATTTTGCCTTAAGGCCACAGAAATGCGAGGACGACCATGTAGTG 420
 DB 361 GGGCTCAGCATAGATTTTGCCTTAAGGCCACAGAAATGCGAGGACGACCATGTAGTG 420

QY 421 TCCATATTGGCACTACCCAGAACAGATTTAAAAATTAACAAGTAACTATATCAGCT 480
 DB 421 TCCATATTGGCACTACCCAGAACAGATTTAAAAATTAACAAGTAACTATATCAGCT 480

QY 481 CGAAGCTATCATGTAATGTTTAAAGAAACATCTATTTAAACACGATCCTCTTAAAAA 540
 DB 481 CGAAGCTATCATGTAATGTTTAAAGAAACATCTATTTAAACACGATCCTCTTAAAAA 540

QY 541 CAAGCATATTTCCAAAAGACAATATATGTTCAGTTTCAACATCTAAGACGACAAA 600
 DB 541 CAAGCATATTTCCAAAAGACAATATATGTTCAGTTTCAACATCTAAGACGACAAA 600

QY 601 TTATATCGAAAGGTAAAGTATGACGTTTCAATTTTCTTTGATTTTGTGTT 660
 DB 601 TTATATCGAAAGGTAAAGTATGACGTTTCAATTTTCTTTGATTTTGTGTT 660

QY 661 ATTTGTTTTTATATACATTTTCTCTTCAATAGATGTTTTTCTCCGATTTTAAAA 720
 DB 661 ATTTGTTTTTATATACATTTTCTCTTCAATAGATGTTTTTCTCCGATTTTAAAA 720

QY 721 ATGACTATAAAGTATTTTATATTAAGACGACGATGCTAGATTCGTTCAAAATC 780
 DB 721 ATGACTATAAAGTATTTTATATTAAGACGACGATGCTAGATTCGTTCAAAATC 780

QY 781 TTTCTGATTTTTTAAAGCTAGTTGGCAACCTGTTTCTTCAAGAAATTTGATTTT 840
 DB 781 TTTCTGATTTTTTAAAGCTAGTTGGCAACCTGTTTCTTCAAGAAATTTGATTTT 840

QY 841 TTCAAAAAAATAGTTTATTTCTCTTATTAATAAGAAACATCTAGAAAAATGAGT 900
 DB 841 TTCAAAAAAATAGTTTATTTCTCTTATTAATAAGAAACATCTAGAAAAATGAGT 900

QY 901 TGCCAGACTAGCCCTAGATGTTTCCCAATAATTAATCATCTGCTGTAATTTTGG 960
 DB 901 TGCCAGACTAGCCCTAGATGTTTCCCAATAATTAATCATCTGCTGTAATTTTGG 960

QY 961 GCCAGCCCCATAATTTTAAACGAAACTGAATCGACGAAACCAATCTGACCTAT 1020
 DB 961 GCCAGCCCCATAATTTTAAACGAAACTGAATCGACGAAACCAATCTGACCTAT 1020

DB 961 GCCAGCCCCATAATTTTAAACGAAACTGAATCGACGAAACCAATCTGACCTAT 1020

QY 1021 TTCTCTAGATTTGTAATAAGGAGAGAGAGAAATCGATTTTAAGTATGTCCC 1080
 DB 1021 TTCTCTAGATTTGTAATAAGGAGAGAGAGAAATCGATTTTAAGTATGTCCC 1080

QY 1081 TGAGATGTGCGGTTTGGCAACGATAGCCCGTAATCATAGCTCATAGGTGCTAGCTCA 1140
 DB 1081 TGAGATGTGCGGTTTGGCAACGATAGCCCGTAATCATAGCTCATAGGTGCTAGCTCA 1140

QY 1141 GGTTCGGACGCTCTCGTGTATCTCATAGGACATATACATGCTTGTTCACCGTTGCTG 1200
 DB 1141 GGTTCGGACGCTCTCGTGTATCTCATAGGACATATACATGCTTGTTCACCGTTGCTG 1200

QY 1201 TTGTTCCATTCGTCACGCTTTCCTTATTTGAAACCAAGAGATACCTATCCCAACAT 1260
 DB 1201 TTGTTCCATTCGTCACGCTTTCCTTATTTGAAACCAAGAGATACCTATCCCAACAT 1260

QY 1261 CCATCTTACTCATGCAACTTCCATGCAAAACGACATATGTTTCTGAAACAGATCTATT 1320
 DB 1261 CCATCTTACTCATGCAACTTCCATGCAAAACGACATATGTTTCTGAAACAGATCTATT 1320

QY 1321 AAAGATCAACAAGCTAGCGTTCTCCGCTAGCTTCCCTCTCTCTGCGATCTTTT 1380
 DB 1321 AAAGATCAACAAGCTAGCGTTCTCCGCTAGCTTCCCTCTCTCTGCGATCTTTT 1380

QY 1381 CGTCCACGACCATG 1394
 DB 1381 CGTCCACGACCATG 1394

RESULT 2
 AAH76333
 ID AAH76333 standard; DNA; 1394 BP.
 AC AAH76333;
 XX
 DT 29-OCT-2001 (first entry)
 XX
 DE Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.
 XX
 KW Ms45; male tissue; regulatory region; transcription; male fertility;
 XX hybrid seed; ds.
 OS Zea mays.
 XX
 PN W020016097-A2.
 PD
 PD 23-AUG-2001.
 PF 13-FEB-2001; 2001WO-US004527.
 PR 15-FEB-2000; 2000US-00504487.
 PA (PION-) PIONEER HI-BRED INT INC.
 PI Albreten MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
 XX WPI, 2001-514772/56.
 DR
 XX
 PT A male tissue-preferred regulatory region comprising nucleotide sequences
 PT essential for initiating transcription of the Ms45 gene useful for
 PT mediating fertility in a male plant.
 XX
 PS Claim 4; Page 47; 50pp; English.
 XX
 CC The invention provides a male tissue-preferred regulatory region (I)
 CC comprising nucleotide sequences essential for initiating transcription of
 CC the Ms45 gene. A method of mediating male fertility in a plant is
 CC provided that involves introducing an expression vector comprising a
 CC promoter operably linked to (I) into a plant where the exogenous gene
 CC impacts male fertility of the plant and (II) controls expression of the
 CC exogenous gene. A method of producing hybrid seeds is also provided. The

CC present sequence represents a nucleic acid sequence encoding an Ms45 male
 CC -tissue preferred regulatory region from Z. mays
 XX
 SQ Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;
 Query Match 100.0%; Score 1394; DB 5; Length 1394;
 Best Local Similarity 100.0%; Pred. No. 2, 2e-291;
 Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAATGCTCTATGAAAAAGATGATACATGTCTATATCCGTTTTCTTAAGGCTC 60
 DB 1 CCAATGCTCTATGAAAAAGATGATACATGTCTATATCCGTTTTCTTAAGGCTC 60
 QY 61 CTTCCTTCTGCTTATTAATGATGAAATCGGGGTTACAAAAAATCTTCCAGGGTGCATGAT 120
 DB 61 CTTCCTTCTGCTTATTAATGATGAAATCGGGGTTACAAAAAATCTTCCAGGGTGCATGAT 120
 QY 121 CTCATGTTTCCACTTCTCCACCTCGGCTTGCACTTCTTGATGTCGGTGGTCCCAT 180
 DB 121 CTCATGTTTCCACTTCTCCACCTCGGCTTGCACTTCTTGATGTCGGTGGTCCCAT 180
 QY 181 CTGACCGAGGCCCATCAGACACCTTTGGGACACCCATCAAGGGCTTTTGGATGGCCCA 240
 DB 181 CTGACCGAGGCCCATCAGACACCTTTGGGACACCCATCAAGGGCTTTTGGATGGCCCA 240
 QY 241 CGAGAGCTATCGGGTCTGGTGAATCCAGGGGATATATGTCCCCCAAAATGTCACCTATA 300
 DB 241 CGAGAGCTATCGGGTCTGGTGAATCCAGGGGATATATGTCCCCCAAAATGTCACCTATA 300
 QY 301 TTATTTATCTTTAGATATATTTTATTTTGAAGAAAAATACAACTTATCTTTTGTGA 360
 DB 301 TTATTTATCTTTAGATATATTTTATTTTGAAGAAAAATACAACTTATCTTTTGTGA 360
 QY 361 GGGGCTTCAGATAGATTTTGGGAGCCAGAAATGCGAGGACGAGCAGCATGTCTATG 420
 DB 361 GGGGCTTCAGATAGATTTTGGGAGCCAGAAATGCGAGGACGAGCAGCATGTCTATG 420
 QY 421 TCCACTTTTGGCACTACCCAGAACAGATTTAAAAAAATAACCAAGTAACTATCCACT 480
 DB 421 TCCACTTTTGGCACTACCCAGAACAGATTTAAAAAAATAACCAAGTAACTATCCACT 480
 QY 481 CGAAGCTATCATGTAATGTTTAAAGAAACATCTATTAACACGATCTCTTTAAAAAA 540
 DB 481 CGAAGCTATCATGTAATGTTTAAAGAAACATCTATTAACACGATCTCTTTAAAAAA 540
 QY 541 CAAAGCATTTTGAAGAGACAAATTAATGTTTACAGTTTACAAACATCTAAGACGACAAA 600
 DB 541 CAAAGCATTTTGAAGAGACAAATTAATGTTTACAGTTTACAAACATCTAAGACGACAAA 600
 QY 601 TTATATCGAAAGGTAAGCTATGACGTTCAAGATTTTCTTTTCAATCTTGTATTTTGT 660
 DB 601 TTATATCGAAAGGTAAGCTATGACGTTCAAGATTTTCTTTTCAATCTTGTATTTTGT 660
 QY 661 ATTGTTTTTATATACATTTTCTTCTTACATAGATGATTTTCTCGAATTTTAA 720
 DB 661 ATTGTTTTTATATACATTTTCTTCTTACATAGATGATTTTCTCGAATTTTAA 720
 QY 721 ATGACTATTAAGTCAATTTTATATAGAGACGAGATGCTAGATTTCTCGTTCAAAATC 780
 DB 721 ATGACTATTAAGTCAATTTTATATAGAGACGAGATGCTAGATTTCTCGTTCAAAATC 780
 QY 781 TTTCTGATTTTTTAAAGCTAGTTTGGCAACCTGTTCTTTTCAAGAAATTTGATTT 840
 DB 781 TTTCTGATTTTTTAAAGCTAGTTTGGCAACCTGTTCTTTTCAAGAAATTTGATTT 840
 QY 841 TTCAAAAAAATAGTTTATTTCTTTTAAATAAGAAACCTTAGAAAAATAGAGT 900
 DB 841 TTCAAAAAAATAGTTTATTTCTTTTAAATAAGAAACCTTAGAAAAATAGAGT 900
 QY 901 TGCCAGACTAGCCCTAGAAATGTTTCCCAATTAATTCATCACTGTGTATTAATTTG 960
 DB 901 TGCCAGACTAGCCCTAGAAATGTTTCCCAATTAATTCATCACTGTGTATTAATTTG 960

QY 961 GCCAGCCCATTAATTTTAAACGGAACGTGAATCGAGCGGAACCAATCTGAGCTAT 1020
 DB 961 GCCAGCCCATTAATTTTAAACGGAACGTGAATCGAGCGGAACCAATCTGAGCTAT 1020
 QY 1021 TTCTTAGATTTAGTAAAG 1080
 DB 1021 TTCTTAGATTTAGTAAAG 1080
 QY 1081 TGAAGATGTCGGTGTGGCAACGATAGCCACCGTATATAGTCTATAGTGTCTAGTCA 1140
 DB 1081 TGAAGATGTCGGTGTGGCAACGATAGCCACCGTATATAGTCTATAGTGTCTAGTCA 1140
 QY 1141 GGTTCGGAGCTCTGTCGATCTCATGAGGATACATGATGCTGTGAACCGTTCGTC 1200
 DB 1141 GGTTCGGAGCTCTGTCGATCTCATGAGGATACATGATGCTGTGAACCGTTCGTC 1200
 QY 1201 TTGTTTCATCTGTCGAAGCCCTTGCTTATGAAACCAAGAGATACCTATCCAAACAT 1260
 DB 1201 TTGTTTCATCTGTCGAAGCCCTTGCTTATGAAACCAAGAGATACCTATCCAAACAT 1260
 QY 1261 CCATCTTACTCATGCACTTCCATGCAACACGACATATGTTTCTGAAACAGATCTATT 1320
 DB 1261 CCATCTTACTCATGCACTTCCATGCAACACGACATATGTTTCTGAAACAGATCTATT 1320
 QY 1321 AAAGATCAGAACAGCTAGGGTCTCCGCTAGCTTCCCTCTCCCTGCGCATCTTTT 1380
 DB 1321 AAAGATCAGAACAGCTAGGGTCTCCGCTAGCTTCCCTCTCCCTGCGCATCTTTT 1380
 QY 1381 CGTCCACACCATG 1394
 DB 1381 CGTCCACACCATG 1394

RESULT 3
 AAX07408
 ID AAX07408 standard; DNA; 1394 BP.
 XX
 AC AAX07408;
 XX
 DT 08-JUN-1999 (first entry)
 XX
 DE Zea mays Ms45 male tissue-preferred regulatory region.
 XX
 KW Ms45; male; tissue-preferred; regulatory region; plant cells;
 KW plant tissue; differentiated; maize; hybrid seed; fertility; ss.
 OS Zea mays.
 OS
 PN W09859061-A1.
 XX
 PD 30-DEC-1998.
 XX
 PF 19-JUN-1998; 98WO-US012895.
 XX
 PR 23-JUN-1997; 97US-00880499.
 XX
 PA (PION-) PIONEER HI-BRED INT INC.
 XX
 PI Albertsen MC, Fox TW, Garnaat CW, Huffman GA, Kendall TL;
 XX
 DR WPI; 1999-105628/09.
 XX
 PT New nucleic acid encoding a Ms45 male tissue-preferred regulatory region
 XX
 PS - useful in mediating plant fertility, especially hybrid seed production.
 XX
 PS Claim 2; Page 22-23; 39pp; English.
 CC The sequence is that encoding an Ms45 male tissue-preferred regulatory
 CC region. It may be used in the construction of a vector for a method of
 CC producing exogenous genes in a male tissue-preferred manner, which is
 CC useful in restoring or conferring fertility, such as in hybrid seed
 CC production. In conferring fertility, a monocot/dicot plant is transformed
 CC with the exogenous nucleotide sequence (a male sterility gene, preferably

CC promoter operably linked to (i) into a plant where the exogenous gene
 CC impacts male fertility of the plant and (ii) controls expression of the
 CC exogenous gene. A method of producing hybrid seeds is also provided. The
 CC present sequence represents a nucleic acid sequence encoding an Ms45 male
 CC -tissue preferred regulatory region from Z. mays
 XX
 SQ Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;

Query Match 99.7%; Score 1389.2; DB 5; Length 1394;
 Best Local Similarity 99.8%; Pred. No. 2,4e-290;
 Matches 1391; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCATGGTGTCTATGAAAAAGATGAGTACATGTGTCTATATCCGTTTCTTAAGGCTCC 60
 DB 1 CCATGGTGTCTATGAAAAAGATGAGTACATGTGTCTATATCCGTTTCTTAAGGCTCC 60
 QY 61 CTTCTTCTGCTTATTAAGTACTGTAATCGGGTTTACAAAAACTTCCACGGGTGATGAT 120
 DB 61 CTTCTTCTGCTTATTAAGTACTGTAATCGGGTTTACAAAAACTTCCACGGGTGATGAT 120
 QY 121 CTGCATGTTCCACTTCTCCCACTCGCGTGGCAATTTCTTGGATGTGGGTGCCAT 180
 DB 121 CTGCATGTTCCACTTCTCCCACTCGCGTGGCAATTTCTTGGATGTGGGTGCCAT 180
 QY 181 CTGACCGAGGCGCATGACACACCTTTCGGGACACCCATCAAGGGCGTTTGGATGGCCCA 240
 DB 181 CTGACCGAGGCGCATGACACACCTTTCGGGACACCCATCAAGGGCGTTTGGATGGCCCA 240
 QY 241 CGAGACGTATCGGGTGTGTGATCCAGGGATATATGTCCCCACATGTGACCTATAT 300
 DB 241 CGAGACGTATCGGGTGTGTGATCCAGGGATATATGTCCCCACATGTGACCTATAT 300
 QY 301 TTATTAATCTTTTGAATTAATTAATTTTGGAAAAATACAACTTAATCTTTTGTGTA 360
 DB 301 TTATTAATCTTTTGAATTAATTAATTTTGGAAAAATACAACTTAATCTTTTGTGTA 360
 QY 361 GGGGCTCAGATGATTTTGGCTTAGGGCCAGAAATGAGAGACGACGATGCTAGTG 420
 DB 361 GGGGCTCAGATGATTTTGGCTTAGGGCCAGAAATGAGAGACGACGATGCTAGTG 420
 QY 421 TCCACTATTGGCACTACCCAGAACAGATTTAAAAAAATACCAAGTAATCAATCCACT 480
 DB 421 TCCACTATTGGCACTACCCAGAACAGATTTAAAAAAATACCAAGTAATCAATCCACT 480
 QY 481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAACACGATCCTTTAAAAAA 540
 DB 481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAACACGATCCTTTAAAAAA 540
 QY 541 CAAGCATATTTGAAAAAGACAAATTAATGTACAGTTTCAACATCTAAGAGCGACAA 600
 DB 541 CAAGCATATTTGAAAAAGACAAATTAATGTACAGTTTCAACATCTAAGAGCGACAA 600
 QY 601 TTATATCGAAAGGTAAGCTATGACGTTCAAGATTTTCTTTTCATCTGTATTTTGT 660
 DB 601 TTATATCGAAAGGTAAGCTATGACGTTCAAGATTTTCTTTTCATCTGTATTTTGT 660
 QY 661 ATTGTTTTTATATACATTTTCTTCTTACATAGAGTGATTTTCTTCGATTTTAA 720
 DB 661 ATTGTTTTTATATACATTTTCTTCTTACATAGAGTGATTTTCTTCGATTTTAA 720
 QY 721 ATGACTATAAGCTATTTTATATAAGAGACGATGTGTATCTGTCTCAAAAATC 780
 DB 721 ATGACTATAAGCTATTTTATATAAGAGACGATGTGTATCTGTCTCAAAAATC 780
 QY 781 TTTCTGATTTTTTAAAGAGTATTTGGCAACCTGTCTTCAAGAATTTTGAATTT 840
 DB 781 TTTCTGATTTTTTAAAGAGTATTTGGCAACCTGTCTTCAAGAATTTTGAATTT 840
 QY 841 TTCAAAAAAATTAAGTTATTTTCTTATTAATAAGAAACACTTGAATAATAGAGT 900
 DB 841 TTCAAAAAAATTAAGTTATTTTCTTATTAATAAGAAACACTTGAATAATAGAGT 900
 QY 901 TGCACAGTACGCTTGAATGTTTCCCAATAAATTAACAATCACTGTGTATATTTTG 960

DB 901 TGCACAGTACGCTTGAATGTTTCCCAATAAATTAACAATCACTGTGTATATTTTG 960
 QY 961 GCCAGCCCATTAATATTATTTAAACCGAAATCGAATCGAGCAACCAATCTGAGCTAT 1020
 DB 961 GCCAGCCCATTAATATTATTTAAACCGAAATCGAATCGAGCAACCAATCTGAGCTAT 1020
 QY 1021 TTCTTGAATTTAGTAAAAAGGAGAGAGAGAGAGAAATCACTTTAAAGTCAATGCTCC 1080
 DB 1021 TTCTTGAATTTAGTAAAAAGGAGAGAGAGAGAGAAATCACTTTAAAGTCAATGCTCC 1080
 QY 1081 TGAGATGTGGGTTTGGCAACGATAGCAACCGTAATCATAGCTCATAGGCTAGCTCA 1140
 DB 1081 TGAGATGTGGGTTTGGCAACGATAGCAACCGTAATCATAGCTCATAGGCTAGCTCA 1140
 QY 1141 GGTTCGGAGCTCTCTGTCTATCTCATATGACATGACATGATGCTTCAACCGTCTGTC 1200
 DB 1141 GGTTCGGAGCTCTCTGTCTATCTCATATGACATGACATGATGCTTCAACCGTCTGTC 1200
 QY 1201 TTGTTTCATGCTCCAAAGCTTGGCTTATTTGAAACCAAGATACCTACTCCAAACAT 1260
 DB 1201 TTGTTTCATGCTCCAAAGCTTGGCTTATTTGAAACCAAGATACCTACTCCAAACAT 1260
 QY 1261 CCATCTTACATGCAACTTCCATGCAACACGACATATGTTTCTGAAACGATCTATT 1320
 DB 1261 CCATCTTACATGCAACTTCCATGCAACACGACATATGTTTCTGAAACGATCTATT 1320
 QY 1321 AAAGATCACAACAGCTAGGCTTCTCCGCTAGCTTCTCTCTCTCCGATCTTTT 1380
 DB 1321 AAAGATCACAACAGCTAGGCTTCTCCGCTAGCTTCTCTCTCTCCGATCTTTT 1380
 QY 1381 CGTCCACGACCATG 1394
 DB 1381 CGTCCACGACCATG 1394

RESULT 5

AAH76340
 ID AAH76340 standard; DNA; 255 BP.

XX AAH76340;

XX 29-OCT-2001 (first entry)

XX Z. mays Ms45 promoter fragment.

XX Ms45; male tissue; regulatory region; transcription; male fertility;

XX hybrid seed; promoter; ds.

XX Zea mays.

XX WO200160997-A2.

XX 23-AUG-2001.

XX 13-FEB-2001; 2001WO-US004527.

XX 15-FEB-2000; 2000US-00504487.

XX (PION-) PIONEER HI-BRED INT INC.

XX Albrethen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;

XX WPI; 2001-514772/56.

XX A male tissue-preferred regulatory region comprising nucleotide sequences
 XX essential for initiating transcription of the Ms45 gene useful for
 XX mediating fertility in a male plant.

XX Example 5; Fig 8; 50pp; English.

XX The invention provides a male tissue-preferred regulatory region (i)
 CC comprising nucleotide sequences essential for initiating transcription of

CC the MS45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (II) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a Z. mays Ms45 promoter fragment
XX
SQ Sequence 255 BP; 59 A; 86 C; 39 G; 71 T; 0 U; 0 Other;
Query Match 17.5%; Score 243.6; DB 5; Length 255;
Best Local Similarity 98.4%; Pred. No. 6e-43;
Matches 246; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1145 CCGCAGCTCTCGTGCATCTCACATGACATACATGCTTTGTTCAACGGTTGCTTGT 1204
DB 5 CCGCGATCCCGTGTATCTCAATGCGATCTCACTGCTTGTCAACGGTTGCTTGT 64
QY 1205 TCCATGCTCCAGCGCTTGCTTATTTGAAACCAAGAGATACCTACTCCAAACATCAT 1264
DB 65 TCCATGCTCCAGCGCTTGCTTATTTGAAACCAAGAGATACCTACTCCAAACATCAT 124
QY 1265 CTACTCATGCAACTTCCATGCAACAGCAGCATATGTTTCCGAAACATCTATTAAAG 1324
DB 125 CTACTCATGCAACTTCCATGCAACAGCAGCATATGTTTCCGAAACATCTATTAAAG 184
QY 1325 ATCACAACAGCTGAGCGTTCCCGTAGCTTCCCTCTCTGCTGCGCATCTTTTCGTC 1384
DB 185 ATCACAACAGCTGAGCGTTCCCGTAGCTTCCCTCTCTGCTGCGCATCTTTTCGTC 244
QY 1385 CACCAACATG 1394
DB 245 CACCAACATG 254
RESULT 6
ID AAH76334 standard; DNA; 158 BP.
XX
XX AAH76334,
AC
XX
DT 29-OCT-2001 (first entry)
XX
DE Z. mays Ms45 male tissue-preferred regulatory region fragment.
XX
XX Ms45; male tissue; regulatory region; transcription; male fertility;
KM hybrid seed; ds.
XX
XX Zea mays.
OS
XX
PN WO200160997-A2.
XX
PD 23-AUG-2001.
XX
PR 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Alberteen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
XX WPI; 2001-514772/56.
DR
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the MS45 gene useful for
PT mediating fertility in a male plant.
PS
XX Claim 5; Page 47; 50pp; English.
CC The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the MS45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a

CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (II) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a DNA fragment -38 to -195 bases upstream of
CC the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
XX
SQ Sequence 158 BP; 41 A; 50 C; 21 G; 46 T; 0 U; 0 Other;
Query Match 10.5%; Score 146; DB 5; Length 158;
Best Local Similarity 99.4%; Pred. No. 6.6e-22;
Matches 157; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1155 CGTGCATCTGCATGAGGATACATGCTTTGTTCAACGGTTGCTTGT 1213
DB 1 CGTGCATCTGCATGAGGATACATGCTTTGTTCAACGGTTGCTTGTTCATTCGTC 60
QY 1214 CAAGCTTGCTTATTTCTGAACCAAGAGATACCTACTCCAAACATCATCTACTCAT 1273
DB 61 CAAGCTTGCTTATTTCTGAACCAAGAGATACCTACTCCAAACATCATCTACTCAT 120
QY 1274 GCACTTCATGCAACACGACATATGTTTCTGAC 1311
DB 121 GCACTTCATGCAACACGACATATGTTTCTGAC 158
RESULT 7
ID AAL15210/c standard; cDNA; 883 BP.
XX
XX AAL15210;
AC
XX
DT 07-DEC-2001 (first entry)
XX
DE Human breast cancer expressed polynucleotide 7667.
XX
XX Human; breast cancer; cell marker; cytostatic; ss.
KM
XX
OS Homo sapiens.
XX
PN WO200151628-A2.
XX
PD 19-JUL-2001.
XX
PF 10-JAN-2001; 2001WO-US000798.
XX
PR 14-JAN-2000; 2000US-0176077P.
PR 14-MAR-2000; 2000US-0189167P.
PR 24-MAR-2000; 2000US-0192099P.
PR 29-MAR-2000; 2000US-0193480P.
PR 15-MAY-2000; 2000US-0205230P.
PR 09-JUN-2000; 2000US-0211315P.
PR 25-JUL-2000; 2000US-0220534P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Lillie J, Xu Y, Wang Y, Steinmann K;
XX
XX WPI; 2001-451856/48.
DR
XX
PT New peptide useful as a marker for the diagnosis of breast cancer.
XX
XX
PS Claim 1; Page 1378; 3695pp; English.
CC The invention relates to human breast cancer expressed polynucleotides
CC (AAL07544-AAL26789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterizing, treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity

XX Sequence 883 BP; 322 A; 18 C; 23 G; 396 T; 0 U; 124 Other;
SQ
Query Match 5.0%; Score 70.2; DB 4; Length 883;
Best Local Similarity 39.2%; Pred. NO. 2.5e-05;
Matches 237; Conservative 0; Mismatches 367; Indels 0; Gaps 0;
QY 419 TGTCCACTATTTGGCACTACCCAGACAGATTTAAAAAATTAACCAAGTAAGTAATCCCA 478
DB 873 TGTNNNNNNNACCCTTAAACCTTTAAANATNTNNNANTTNNANNNAAAAAATTTTAAANN 814
QY 479 CTCGAAAGCTATCATGTATGTGTTTAAAGAAACATCTTAAAAACGATCCTCTTAAAA 538
DB 813 TAAAAATTTTTTAAAAAATTTTAAATNTNANTNANTTTTTNTTAAAAAANNNAAAAA 754
QY 539 AACAGACATTTTCGAAAGACAAATTTAGTTACAGTTTACAACATCTTAAAGCGACA 598
DB 753 TTAANNNTTNTTAAATTTATTTAAACCAAAATTTTTTTAAAAAATTTTTTAAANNTT 694
QY 599 AATTATATCGAAAGCTAGCTATGAGCTTGCAGATTTTCTTTTCATCTTGTATTTTG 658
DB 693 AATTAATTTTAAATTTTTTTTAAATNNAAAAAATTTTAAATTTTAAACAAATNTTTTTT 634
QY 659 TTAATGTTTTTATATACATTTCTCTCTACATAGAGATTTTCTTCGATTTTATA 718
DB 633 TTTNTTTNNATTAATAAATTTTAAATTAATAAANNNANNTTTTTTTATNAAATATTTAA 574
QY 719 AAATGACTAATAAGCTATTTTAAATTAAGAGACGAGATGCTAGATTCGTTCAAAA 778
DB 573 AAAAAAANNTTTTTTTTTTATANTTTATTAATAAATTTTATNTNTTTTTCTTNAANA 514
QY 779 TCTTTCTGATTTTTTTAAGAGCTAGTTGGCAACCTGTTCTTCAAGATTTTGAT 838
DB 513 AAAAAAATTTTTTAAANNTTTTTTAAACCTTNAATNAAANANAATTTTTTNNNTT 454
QY 839 TTTTCAAAAAATTTAGTTATTTCTCTTATTAATAAGAAACCTTGAATAATAGA 898
DB 453 NNNANNTTAAATAATTTTATTTTNTTANTCTATTAATTAATAAANNNANNTAATAAT 394
QY 899 GTTGCCAGACTAGCCCTAGAAATGTTTCCCAATTAATTAACATCACTGTGTATATAT 958
DB 393 TTTNAAAAAATTAATTTATANTTATNNAANNAANAATTTAAANNTNANNNAAAAATTTTA 334
QY 959 TGGCCAGCCCATTAATTTATTAACCGAAACTGAAATGAGCGAAACCAATCTGAGCT 1018
DB 333 NTTTAAATTAATAAATAATTAATTTAAACNTCAATTTNTAATTTTAAATTAATAAATAAATAAT 274
QY 1019 ATTT 1022
DB 273 NTTT 270
RESULT 8
ACN85231/c
ID ACN85231 standard; DNA; 960 BP.
XX ACN85231;
XX
XX 02-DEC-2004 (first entry)
XX
XX Breast cancer related marker, seq id 6381.
XX
XX Cancer; breast; tumour; cytostatic; marker; detection; therapy; ds.
XX
XX Homo sapiens.
XX OS
XX PN US2003099974-A1.
XX
XX 29-MAY-2003.
XX PD
XX 18-JUL-2002; 2002US-00198846.
XX PF
XX 18-JUL-2001; 2001US-0306220P.
XX

XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Lillie J, Xu Y, Wang Y, Steinmann K;
XX
DR WPI; 2003-787014/74.
XX
PT Novel isolated polypeptide associated with breast cancer, useful for
PT detecting presence of polypeptide in sample, as a marker for breast
PT cancer.
XX
XX Disclosure; SEQ ID NO 6381; 36pp; English.
XX
XX The invention relates to an isolated polypeptide (I) associated with
CC breast cancer which is encoded by a nucleic acid molecule comprising a
CC nucleotide sequence (SI). Further disclosed is an antibody that binds to
CC the polypeptide of the invention. The activity of the polypeptide of the
CC invention may be described as cytostatic. The antibody is useful for
CC detecting the presence of (I) in a sample. Nucleic acid molecules of the
CC invention are useful in the detection of breast tumours. (II) is useful as
CC a marker for breast cancer and in breast cancer therapy. Sequences given
CC in records ACN78851-ACN92934 represent nucleic acid markers associated
CC with breast cancer. Note: The sequence listing does not form part of the
CC specification but may be obtained in electronic format from the USPTO web
CC site at seqdata.uspto.gov/sequence.html?docID=2003099974
XX
SQ Sequence 960 BP; 340 A; 33 C; 39 G; 421 T; 0 U; 127 Other;
Query Match 5.0%; Score 70.2; DB 11; Length 960;
Best Local Similarity 39.2%; Pred. NO. 2.5e-05;
Matches 237; Conservative 0; Mismatches 367; Indels 0; Gaps 0;
QY 419 TGTCCACTATTTGGCACTACCCAGACAGATTTAAAAAATTAACCAAGTAAGTAATCCCA 478
DB 933 TGTNNNNNNNACCCTTAAACCTTTAAANATNTNNNANTTNNANNNAAAAAATTTTAAANN 874
QY 479 CTCGAAAGCTATCATGTATGTGTTTAAAGAAACATCTATTAACCGATCCTCTTAAAA 538
DB 873 TAAAAATTTTTTAAAAAATTTTAAATNTNAAATTTTNTTAAAAAANNNAAAAA 814
QY 539 AACAGACATTTTCGAAAGACAAATTTAGTTACAGTTTACAACATCTAAGAGCGACA 598
DB 813 TTAANNTTTNTTAAATTTTAAACCAAAATTTTTTTTAAAAAATTTTTTAAANNT 754
QY 599 AATTAATTCGAAAGCTAGCTATGAGCTTGCAGATTTTCTTTTCATCTTGTATTTTG 658
DB 753 AATTAATTTAAATTTTTTTTAAATTAATAAATAAATAATTTTAAACAAATNTTTTTT 694
QY 659 TTAATGTTTTTATATACATTTCTCTCTACATAGAGCTGATTTCTCGATTTTATA 718
DB 693 TTTNTTTNNATTAATAAATAATTTTAAATTAATAAANNNANNTTTTTTTTAAANNTATTTTA 634
QY 719 AAATGACTAATAAGCTATTTTAAATTAAGAGACGAGATGCTAGATTCGTTCAAAA 778
DB 633 AAAAAAANNTTTTTTTTTTATANTTTATTAATAAATAATTTATNTNTTTTTCTTNAANA 574
QY 779 TCTTTCTGATTTTTTTAAGAGCTAGTTGGCAACCTGTTCTTCAAGATTTTGAT 838
DB 573 AAAAAAATTTTTTAAANNTTTTTTAAACCTTNAATNAAANANAATTTTTTNNNTT 514
QY 839 TTTTCAAAAAATTTAGTTATTTCTCTTATTAATAAGAAACCTTGAATAATAGA 898
DB 513 NNNANNTTAAATAATTTATTTTNTTANTCTATTAATTAATAAANNNANNTAATAAT 454
QY 899 GTTGCCAGACTAGCCCTAGAAATGTTTCCCAATTAATTAACATCACTGTGTATATAT 958
DB 453 TTTNAAAAAATTAATTTTATANTTATNNAANNAANAATTTTAAANNTNANNNAAAAATTTTA 394
QY 959 TGGCCAGCCCATTAATTTATTAACCGAAACTGAAATGAGCGAAACCAATCTGAGCT 1018
DB 393 NTTTAAATTAATAAATAATTAATTTAAACNTCAATTTNTAATTTTAAATTAATAAATAAATAAT 334
QY 1019 ATTT 1022

Db 333 NTTT 330

RESULT 9
ADRo4296
ID ADRo4296 strand; DNA; 13400 BP.
XX
AC ADRo4296;
XX
DT 04-NOV-2004 (first entry)
XX
DE Corn FT homologue nucleotide sequence SEQ ID NO:63.
XX
KM flowering locus T gene; FT; terminal flower; TFL; Apetala3; Ap3; plant;
KM floral development; plant sterility; plant fertility; flowering time;
KM plant growth rate; inflorescence architecture; tissue culture morphology;
KM cell division; FT homologue; gene; ds.
XX
OS Zea mays.
XX
PN MO2004067723-A2.
XX
PD 12-AUG-2004.
XX
PF 29-JAN-2004; 2004MO-US002422.
XX
PR 30-JAN-2003; 2003US-00343477.
XX
PA (PION-) PIONEER HI-BRED INT INC.
PA (DUPO) DU PONT DE NEMOURS & CO E I.
PI Danilevskaya O, Hermon P, Bruggemann E, Shitproun D, Ananiev E;
PI Rafalski JA, Sakai H, Cahoon E, Cahoon R, Klein T;
XX
XX MPI; 2004-580996/56.
XX
DR
XX
PT New polynucleotides, specifically nucleic acid fragments encoding
PT flowering locus T gene (FT) or terminal flower (TFL), or Apetala3 (Ap3)
PT homologue, useful for floral development, e.g. engineering plant flowering
PT time.
XX
XX Claim 6; SEQ ID NO 63; 109pp; English.
XX
PS The present invention describes an isolated polynucleotide comprising a
CC first, second, third, fourth or fifth nucleotide sequence, or their
CC complement encoding a polypeptide either having flowering locus T gene
CC (FT), terminal flower (TFL), or Apetala3 (Ap3) homologue activity. Also
CC described: (1) a vector comprising the polynucleotide; (2) a recombinant
CC DNA construct comprising the polynucleotide; (3) transforming a cell by
CC transforming a cell with the polynucleotide; (4) a cell comprising the
CC recombinant DNA construct; (5) producing a plant comprising transforming
CC a plant cell with the polynucleotide, and regenerating a plant from the
CC transformed plant cell; (6) a plant comprising the recombinant DNA
CC construct; (7) a seed comprising the recombinant DNA construct; (8) an
CC isolated polynucleotide comprising a first nucleotide sequence, where the
CC first nucleotide sequence contains at least 30 nucleotides, and where the
CC first nucleotide sequence is comprised by another polynucleotide, where
CC the other polynucleotide includes the second, third, fourth, fifth or
CC sixth nucleotide sequence; (9) an isolated polypeptide having FT or Ap3
CC homologue activity, as described above; and (10) isolating a polypeptide
CC encoded by the polynucleotide comprising isolating the polypeptide from a
CC cell containing a recombinant DNA construct comprising the polynucleotide
CC operably linked to a regulatory sequence. The polynucleotides are useful
CC for floral development, e.g. engineering plant sterility/fertility,
CC flowering time, plant growth rate, inflorescence architecture, and tissue
CC culture morphology and the rate of cell division to enhance
CC transformation. The present sequence represents an FT homologue
CC nucleotide sequence from the present invention.
XX
SQ Sequence 13400 BP, 3962 A, 2948 C, 2695 G, 3795 T, 0 U, 0 Other;

Query Match 5.0%; Score 69.8; DB 13; Length 13400;

Best Local Similarity 73.6%; Pred. No. 6.1e-05;
Matches 89; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 811 ACCGTTCTCTTCAAGATTTTCAATTTTCAAAAAAATTTAGTTATTTCTCTTCA 870
DB 8559 ACCATTTTATTTTCAAGAGGTTTATTTATCAAGAAAAATTTAGTTCATTTCTTGG 8618
QY 871 TAAATAGAAAAACACTTGAATAATGAGTCCGACACGACGCAATGTTTCCCAA 930
DB 8619 AAAAAATAAAAATCAATGAAAAAATGGGTGTCAACTAGTCTTATTAGTTTTCAT 8678
QY 931 T 931
DB 8679 T 8679

RESULT 10
AA210551
ID AA210551 strand; DNA; 2657 BP.

AC AA210551;
XX
DT 16-NOV-1999 (first entry)
XX
DE DNA sequence of the P-Zeap promoter of maize.

KM P-Zeap promoter; male sterile plant; glyphosate tolerance; glyphosate;
KM male reproductive tissue; hybrid seed production; crop outcrossing;
KM flower life; ss.

OS Zea mays.
XX
PN MO9946396-A2.
XX
PD 16-SEP-1999.
XX

PF 09-MAR-1999; 99WO-US005126.
XX
PR 09-MAR-1998; 98US-0077277P.
XX

PA (MONS) MONSANTO CO.

PI Brown SM, Fromm ME;

XX
XX MPI; 1999-551420/46.
XX

PT Production of male sterile plants using a gene encoding glyphosate
PT tolerance, used for, e.g. production of hybrid seed.

PS Disclosure; Fig 1A-B; 54pp; English.

CC The present sequence represents the P-Zeap promoter of maize. It is used
CC in the method of the invention. The specification describes a method for
CC the production of male sterile plants. The method comprises selective
CC expression of DNA encoding a protein that causes tolerance to glyphosate
CC and application of glyphosate. The method uses two DNA molecules, each
CC operably linked to a separate promoter, whereby the first promoter
CC functions in plant cells to produce a protein that causes tolerance to
CC glyphosate, and the second promoter functions in plant cells to cause the
CC expression of a second RNA sequence in a male reproductive tissue.
CC Expression of the second RNA sequence in a male reproductive tissue
CC in which it is expressed. Expression of the second DNA molecule causes
CC the production of an RNA sequence which can inhibit the glyphosate
CC tolerance generated by expression of the first DNA molecule. By using a
CC promoter for the second DNA molecule which restricts the production of
CC the antisense RNA to only a subset of the tissues which express the first
CC DNA molecule, only the subset of tissues in which the second DNA molecule
CC is expressed will be susceptible to glyphosate toxicity. In this way, a
CC specific cell type or combination of cell types, depending upon the
CC promoters utilized, can be selectively ablated by application of
CC glyphosate to the plant. The methods can be used for producing male-
CC sterile plants for use in the production of hybrid seed, for minimizing
CC undesirable crop outcrossing, and for lengthening flower life. The


```
XX Byatt JC, Mathalagan N, Tao N, Warren WC;
PI WPI; 2003-110599/10.
XX New nucleic acid associated with lactation, and muscle and fat
XX deposition, useful for genome mapping, gene identification and analysis,
PT cattle breeding, or for genetically improving cattle.
XX
XX Claim 2; SEQ ID NO 1009; 245pp; English.
XX
XX The invention relates to a purified nucleic acid molecule associated with
XX lactation or muscle and fat deposition (designated LMPD), derived from
XX cattle, and the LMPD nucleic acid can specifically hybridise to a second
XX nucleic acid molecule comprising any of 15112 nucleotide sequences,
XX appearing as ABX34836-ABX49947, or complements of them. Also included are
XX ; (1) a transformed cell having a nucleic acid comprising an LMPD nucleic
XX acid linked to a promoter and a 3' non-translated sequence that
XX functions in the cell to cause termination of transcription and addition
XX of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
XX (2) determining a level or pattern of a molecule in a bovine cell or
XX tissue comprising: (a) incubating a marker nucleic acid (comprising any
XX of the 15112 nucleic acid sequences or its complement or fragment) with a
XX complementary nucleic acid molecule obtained from the bovine cell or
XX tissue, where hybridisation between the marker nucleic acid and the
XX complementary nucleic acid permits the detection of the molecule; and (b)
XX detecting the level or pattern of the complementary nucleic acid, where
XX the detection of the complementary nucleic acid is predictive of the
XX level or pattern of the molecule. The LMPD nucleic acid is used for
XX determining a level or pattern of a molecule in a bovine cell or tissue.
XX It is useful for genome mapping, gene identification and analysis, cattle
XX breeding, preparation of constructs for use in cattle gene expression, or
XX for genetically improving cattle. The present sequence is one of the
XX 15112 bovine LMPD EST (expressed sequence tag) nucleic acids. Note: The
XX present sequence was not shown in the specification but was obtained in
XX CC electronic format from the USPTO web site:
XX seqdata.uspto.gov/sequence.html?docid=20020137139
XX
XX Sequence 439 BP; 45 A; 51 C; 56 G; 286 T; 0 U; 1 Other;
SQ
XX
XX Query Match 4.7%; Score 66.2; DB 8; Length 439;
XX Best Local Similarity 54.3%; Pred. No. 0.00015;
XX Matches 134; Conservative 0; Mismatches 113; Indels 0; Gaps 0;
XX
XX 626 TTCAGATTTCTTTTCATCTCTGTTATTTGTTATGTTTATATACATTTCTCT 665
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 182 TTCGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTTATTTATTTCTT 241
XX
XX 686 CTTACAAATAGAGATTTTCTCGATTTTATTAATGATATTAAGTCATTTTATATA 745
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 242 TTTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 301
XX
XX 746 AGACACGACATCTGTGATTTCTGTTCAAAAATCTTGTGATTTTTTAAAGACTAGTT 805
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 302 TGTGGGGGGGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 361
XX
XX 806 TGGCAACCTGTTCTTTCAAGAAATTTGATTTTCAAAAAAATAGTTATTTCT 865
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 362 TTTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTATTTATTTTATTTT 421
XX
XX 866 CTTTATA 872
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 422 TTCTTTA 428
XX
XX RESULT 15
XX AA187279/c
XX AA187279 standard; cDNA; 346 BP.
XX
XX AA187279;
XX
XX 06-NOV-2001 (first entry)
XX
```

```
DE Human polynucleotide SEQ ID NO 7339.
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX tissue growth factor; immunomodulatory; cancer; leukaemia;
XX nervous system disorders; arthritis; inflammation; ss.
XX
XX Homo sapiens.
XX
XX WO200164835-A2.
XX
XX 07-SEP-2001.
XX
XX 26-FEB-2001; 2001MO-US004927.
XX
XX 28-FEB-2000; 2000US-00515126.
XX
XX 18-MAY-2000; 2000US-00577409.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-514838/56.
XX
XX P-PADB; AA007348.
XX
XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
XX and treating e.g. leukemia, inflammation and immune disorders.
XX
XX Claim 1; SEQ ID NO 7339; 1399pp + Sequence listing; English.
XX
XX The invention relates to human polynucleotides (AA179941-AA193841) and
XX the encoded proteins (AA000010-AA013910) that exhibit activity elating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and
XX activin/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX inflammation. Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic format
XX directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 346 BP; 188 A; 22 C; 16 G; 120 T; 0 U; 0 Other;
SQ
XX
XX Query Match 4.7%; Score 65.6; DB 4; Length 346;
XX Best Local Similarity 52.6%; Pred. No. 0.00019;
XX Matches 143; Conservative 0; Mismatches 129; Indels 0; Gaps 0;
XX
XX 446 AGATTTAAAAAATACCAAGTAATCACTCGAAGAGCTATCATGTAAATGTTTAA 505
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 334 ACAATATAGATATATATTTAAATTAAGAGTTAAATTTATTCATTAATAAATTTTA 275
XX
XX 506 GAACATCTATTTAAACCGCATCTCTTAAAAAACAAGACATTTTCGAAAGACAAAT 565
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 274 GACTTATATATATATATATATATATATATATATATATATATATATATATAT 215
XX
XX 566 TATGTACAGTTTACCAACATCTAAGAGCAATATATATGAAAGTAAAGTATAGAG 625
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 214 TTTTAAATATATATATATATATATATATATATATATATATATATATATATAT 155
XX
XX 626 TTCAGATTTCTTTTCATCTGTGTTATTTGTTATGTTTATATACATTTCTCT 685
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 154 TATTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 95
XX
XX 686 CTTACAAATAGAGATTTTCTCGATTTTAT 717
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 94 TTTTATTTTTTTTTTTTTTTTTTTTTTTTTTT 63
XX
XX Search completed: September 15, 2005, 21:33:19
XX Job time : 676.892 secs
XX
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 14, 2005, 19:14:19 ; Search time 210.859 Seconds
(without alignments)
10817.505 Million cell updates/sec

Title: US-10-713-381-2

Perfect score: 1394

Sequence: 1 cccatgctgcctcatgaaaa.....cttttcgccaccaccatg 1394

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued Patents NA: *
2: /cgn2_6/ptodata/1/lna/5A COMB.seq: *
3: /cgn2_6/ptodata/1/lna/5B COMB.seq: *
4: /cgn2_6/ptodata/1/lna/6A COMB.seq: *
5: /cgn2_6/ptodata/1/lna/6B COMB.seq: *
6: /cgn2_6/ptodata/1/lna/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1394	100.0	1394	3	US-08-880-499-2
2	1389.2	99.7	1394	3	US-08-880-499-1
3	66.4	4.8	6027	2	US-08-968-542C-1
4	66.4	4.8	6027	4	US-09-554-467A-1
5	58.6	4.2	2523	2	US-08-410-784A-3
6	57	4.1	19124	2	US-08-487-826B-13
7	55.8	4.0	279	4	US-09-313-294A-5397
8	55.6	4.0	612	4	US-09-902-540-1357
9	55.6	4.0	1394	3	US-08-880-499-1
10	55.6	4.0	1394	3	US-08-880-499-2
11	55.4	4.0	2614	4	US-09-806-708B-23
12	55.2	4.0	1055	4	US-09-004-056-1
13	55	3.9	396	4	US-09-640-173-53
14	55	3.9	396	4	US-09-713-550-53
15	55	3.9	396	4	US-09-825-294-53
16	55	3.9	396	4	US-09-970-566-53
17	54.6	3.9	55886	4	US-09-949-016-15129
18	53.8	3.9	307	4	US-09-313-294A-4753
19	53.8	3.9	6027	4	US-08-968-542C-1
20	53.8	3.9	6027	4	US-09-554-467A-1
21	52.2	3.7	1141	4	US-09-806-708B-22
22	51.6	3.7	134987	4	US-09-949-016-15348
23	51.6	3.7	134987	4	US-09-949-016-15349
24	51.6	3.7	134987	4	US-09-949-016-15350
25	51.6	3.7	134987	4	US-09-949-016-15507
26	51.6	3.7	134987	4	US-09-949-016-15508
27	51.6	3.7	134987	4	US-09-949-016-15509

C 28	51.4	3.7	1141	4	US-09-806-708B-22	Sequence 22, Appl
C 29	51.4	3.7	50383	4	US-09-949-016-17600	Sequence 17600, A
C 30	51.4	3.7	129415	4	US-09-949-016-16397	Sequence 16397, A
C 31	51	3.7	601	4	US-09-949-016-156535	Sequence 156535, A
C 32	50.6	3.6	1039	4	US-09-902-540-1280	Sequence 1280, Ap
C 33	50.6	3.6	16573	4	US-09-949-016-14876	Sequence 14876, A
C 34	50.6	3.6	18773	4	US-09-949-016-14164	Sequence 14164, A
C 35	50.4	3.6	731	1	US-08-451-405A-2	Sequence 2, Appl1
C 36	50.2	3.6	832	1	US-09-621-976-2813	Sequence 2813, Ap
C 37	50.2	3.6	317366	4	US-09-949-016-16001	Sequence 16001, A
C 38	50	3.6	1392	4	US-09-257-584-1	Sequence 1, Appl1
C 39	49.8	3.6	2435	3	US-09-306-593-1	Sequence 1, Appl1
C 40	49.8	3.6	231129	4	US-09-949-016-16110	Sequence 16110, A
C 41	49.8	3.6	266283	4	US-09-949-016-11934	Sequence 11934, A
C 42	49.6	3.6	440681	4	US-09-790-988-1	Sequence 1, Appl1
C 43	49.2	3.5	126176	4	US-09-949-016-16137	Sequence 16137, A
C 44	49.2	3.5	126176	4	US-09-949-016-16138	Sequence 16138, A
C 45	47.8	3.4	601	4	US-09-949-016-25787	Sequence 25787, A

ALIGNMENTS

RESULT 1
US-08-880-499-2

Sequence 2, Application US/08880499

Patent No. 6037523

GENERAL INFORMATION:

APPLICANT: Alberson, Marc C.

APPLICANT: Fox, Tim W.

APPLICANT: Hufman, Gary A.

APPLICANT: Kendall, Timmy L.

TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION

NUMBER OF SEQUENCES: 2

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.

STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.

STREET: Box 1000

CITY: Johnston

STATE: Iowa

COUNTRY: USA

ZIP: 50131

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/880,499

FILING DATE: CONCURRENTLY HEREWITH

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Sweeney, Patricia A.

REGISTRATION NUMBER: 32,733

REFERENCE/DOCKET NUMBER: 0578

TELECOMMUNICATION INFORMATION:

TELEPHONE: (515) 248-4800

TELEFAX: (515) 248-4844

INFORMATION FOR SEQ. ID NO. 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1394 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-880-499-2

Query Match 100.0% ; Score 1394 ; DB 3 ; Length 1394 ;
Best Local Similarity 100.0% ; Pred. No. 0 ;
Matches 1394 ; Conservative 0 ; Mismatches 0 ; Indels 0 ; Gaps 0 ;

QY 1 CCATGCTGCTCTATGAAAAAGATGATCAATGTCCTATATGCCGTTTCTTAGGGTCC 60
Db 1 CCAATGGTCTCTATGAAAAAGATGATCAATGTCCTATATGCCGTTTCTTAGGGTCC 60
QY 61 CTTCTTCTGCTTATATCTAGTCAATCGGGGTTTACAAAAAATTCCACGGGTGCATGAT 120
Db 61 CTTCTTCTGCTTATATCTAGTCAATCGGGGTTTACAAAAAATTCCACGGGTGCATGAT 120
QY 121 CTCGATGTTCACTTCTCCACCTCGGCTTGCACTTCTTGGAATGTCGGTGGTTCCAT 180
Db 121 CTCGATGTTCACTTCTCCACCTCGGCTTGCACTTCTTGGAATGTCGGTGGTTCCAT 180
QY 181 CTGACCGAGGCCCTCAAGACCTTTGGGACACCCATCAAGGGCTTTTGGATGGGCCA 240
Db 181 CTGACCGAGGCCCTCAAGACCTTTGGGACACCCATCAAGGGCTTTTGGATGGGCCA 240
QY 241 CGAGACGTATCGGGTCTGGTGAATCGAGGGATATATGTCGCCCAATGTCACCTATA 300
Db 241 CGAGACGTATCGGGTCTGGTGAATCGAGGGATATATGTCGCCCAATGTCACCTATA 300
QY 301 TTATATCTCTTATGATATTTATTTTGAATAAATACTTATACTTTTGTGTA 360
Db 301 TTATATCTCTTATGATATTTATTTTGAATAAATACTTATACTTTTGTGTA 360
QY 361 GGGCTTCAGATGATTTTGGCTTAGGGCCCAAAATCGCAGACCGCATGTCATGTC 420
Db 361 GGGCTTCAGATGATTTTGGCTTAGGGCCCAAAATCGCAGACCGCATGTCATGTC 420
QY 421 TCCACTATGGCACTACCCAGAACAGATTTAAAAAATAATACCAAGTATATCACT 480
Db 421 TCCACTATGGCACTACCCAGAACAGATTTAAAAAATAATACCAAGTATATCACT 480
QY 481 CGAAAGCTATCATGTAAATGTTTAAAGAAACATCTATTTAAACACGATCCTTTAAAA 540
Db 481 CGAAAGCTATCATGTAAATGTTTAAAGAAACATCTATTTAAACACGATCCTTTAAAA 540
QY 541 CAAGCATATTTCCAAAAGACAAATTTGTTACGTTTACAAACTTAAAGAGACAA 600
Db 541 CAAGCATATTTCCAAAAGACAAATTTGTTACGTTTACAAACTTAAAGAGACAA 600
QY 601 TTATATGAAAGGATGATGACGTGAGATTTTCTTTTCAATCTGTTATTTTGT 660
Db 601 TTATATGAAAGGATGATGACGTGAGATTTTCTTTTCAATCTGTTATTTTGT 660
QY 661 ATTGTTTATATATATTTCTTCTTACAAATGAGTATTTTCTCCGATTTTAAAA 720
Db 661 ATTGTTTATATATATTTCTTCTTACAAATGAGTATTTTCTCCGATTTTAAAA 720
QY 721 ATGACTATTAAGTCAATTTTATATAGAGACCGCATGTCGTAATCTCGTTCAAAATC 780
Db 721 ATGACTATTAAGTCAATTTTATATAGAGACCGCATGTCGTAATCTCGTTCAAAATC 780
QY 781 TTTCTGATTTTAAAGAGTAGTTTGGCAACCTGTTCTTTCAAAAGATTTGATTT 840
Db 781 TTTCTGATTTTAAAGAGTAGTTTGGCAACCTGTTCTTTCAAAAGATTTGATTT 840
QY 841 TTCAAAAAAATTAGTTATTTTCTCTTTTAAATGAAACACTTAGAAAAATAGAGT 900
Db 841 TTCAAAAAAATTAGTTATTTTCTCTTTTAAATGAAACACTTAGAAAAATAGAGT 900
QY 901 TGGCAGACTGACCTAGAAATGTTTCCCAATTAATTCATCTGTAATTAATTTG 960
Db 901 TGGCAGACTGACCTAGAAATGTTTCCCAATTAATTCATCTGTAATTAATTTG 960
QY 961 GCGAGCCCAATTAATTTTAAACCGAAATGAAATGAGGAAACCAATCTGAGCTAT 1020
Db 961 GCGAGCCCAATTAATTTTAAACCGAAATGAAATGAGGAAACCAATCTGAGCTAT 1020
QY 1021 TTCTCTAGATTAGTAAAGAGAGAGAGAGAGAGAAATCAGTTTAACTCATTTGCC 1080
Db 1021 TTCTCTAGATTAGTAAAGAGAGAGAGAGAGAGAAATCAGTTTAACTCATTTGCC 1080

QY 1081 TGAGATGTCGGCTTTGGCAACGATAGCAACCGTAATCATAGTCAATGAGTCTAGCTCA 1140
Db 1081 TGAGATGTCGGCTTTGGCAACGATAGCAACCGTAATCATAGTCAATGAGTCTAGCTCA 1140
QY 1141 GGTTCGGAGCTCTGCTGTCATCTCATATGAGCATATCACTAGTCTTCAACGCTTCGTC 1200
Db 1141 GGTTCGGAGCTCTGCTGTCATCTCATATGAGCATATCACTAGTCTTCAACGCTTCGTC 1200
QY 1201 TTGTTCCATGTCGAAGCTTGCCTATTTCTGAACCAAGAGATACCTACCTCCAAACAT 1260
Db 1201 TTGTTCCATGTCGAAGCTTGCCTATTTCTGAACCAAGAGATACCTACCTCCAAACAT 1260
QY 1261 CGATCTTACTCAGCACTTCCATGCAACACCGCATATGTTTCTGGAACATATCTAT 1320
Db 1261 CGATCTTACTCAGCACTTCCATGCAACACCGCATATGTTTCTGGAACATATCTAT 1320
QY 1321 AAGATCACAACAGCTAGGCTTCTCCGCTAGGCTTCCCTCTCTGCGCATCTTTT 1380
Db 1321 AAGATCACAACAGCTAGGCTTCTCCGCTAGGCTTCCCTCTCTGCGCATCTTTT 1380
QY 1381 CGTCCACCAACCATG 1394
Db 1381 CGTCCACCAACCATG 1394

RESULT 3
US-08-968-542C-1/C
; Sequence 1, Application US/08968542C
; Patent No. 5981728
; GENERAL INFORMATION:
; APPLICANT: Myers, et al.
; TITLE OF INVENTION: dnull Codes For A No. 5981728el Search
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: McGregor & Adler, LLP
; STREET: 8011 Candle Lane
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word 6.0.1 for Macintosh
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/968,542C
; FILING DATE: No. 5981728ember 12, 1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benjamin Aaron Adler, Ph.D., J.D.
; REGISTRATION NUMBER: 35,423
; REFERENCE/DOCKET NUMBER: D6036
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713) 777-2321
; TELEFAX: (713) 777-6908
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6027 bp
; TYPE: nucleic acid
; STRANDEDNESS: double-stranded
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: cDNA to mRNA
; HYPOTHETICAL: no
; ANTI-SENSE: no
; ORIGINAL SOURCE:
; ORGANISM: maize
; TISSUE TYPE: endosperm

```
/ IMMEDIATE SOURCE:
/ LIBRARY: maize endosperm cDNA library in
/ CLONE: pMgf10; pMgf6a; pMgf6-2M
/ US-08-968-542C-1

Query Match      4.8%; Score 66.4; DB 2; Length 6027;
Best Local Similarity 77.5%; Pred. No. 2.2e-06;
Matches 93; Conservative 0; Mismatches 26; Indels 1; Gaps 1;

QY 792 TTAAAGAGCTAGTTGGCAACCTGTTCTTTCAAGATTGATTTTCAAAAAA 851
DB 5606 TCTACGGGGCTAGTTGGGAACCCCATTT-TTCCAGGGGATTTCCATTTTCCAGAAAA 5548
DB 5547 TTAGTTATTTTTCATTGGAAAAATGAAATCTCTTGAAAAATAGATTCACTACTAG 5488

RESULT 4
US-09-554-467A-1/C
/ Sequence 1, Application US/09554467A
/ Patent No. 6639125
/ GENERAL INFORMATION:
/ APPLICANT: Myers, Alan M.
/ APPLICANT: James, Martha G.
/ TITLE OF INVENTION: dulli Coding for a No. 6639125el Starch Synthase and Uses
/ TITLE OF INVENTION: Thereof
/ FILE REFERENCE: D6036PCT
/ CURRENT APPLICATION NUMBER: US/09/554,467A
/ CURRENT FILING DATE: 2000-05-12
/ PRIOR APPLICATION NUMBER: PCT/US98/24225
/ PRIOR FILING DATE: 1998-11-12
/ PRIOR APPLICATION NUMBER: US 08/062,102
/ PRIOR FILING DATE: 1997-11-12
/ NUMBER OF SEQ ID NOS: 37
/ SEQ ID NO 1
/ LENGTH: 6027
/ TYPE: DNA
/ ORGANISM: maize
/ FEATURE:
/ OTHER INFORMATION: cDNA sequence corresponding to the gene encoding the
/ OTHER INFORMATION: starch synthase enzyme DUL1.
/ US-09-554-467A-1

Query Match      4.8%; Score 66.4; DB 4; Length 6027;
Best Local Similarity 77.5%; Pred. No. 2.2e-06;
Matches 93; Conservative 0; Mismatches 26; Indels 1; Gaps 1;

QY 792 TTAAAGAGCTAGTTGGCAACCTGTTCTTTCAAGATTGATTTTCAAAAAA 851
DB 5606 TCTACGGGGCTAGTTGGGAACCCCATTT-TTCCAGGGGATTTCCATTTTCCAGAAAA 5548
DB 5547 TTAGTTATTTTTCATTGGAAAAATGAAATCTCTTGAAAAATAGATTCACTACTAG 5488

RESULT 5
US-08-410-784A-3/C
/ Sequence 3, Application US/08410784A
/ Patent No. 5912413
/ GENERAL INFORMATION:
/ APPLICANT: MYERS, ALAN M.
/ APPLICANT: JAMES, MARTHA G.
/ TITLE OF INVENTION: ISOLATION OF SUL1, A STARCH DEBRANCHING
/ TITLE OF INVENTION: ENZYME, THE PRODUCT OF THE MAIZE GENE
/ NUMBER OF SEQUENCES: 9
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Weingarten, Schurgin, Gagnebin and Hayes LLP
/ STREET: Ten Post Office Square
/ CITY: Boston
```

```
/ STATE: MA
/ COUNTRY: USA
/ ZIP: 02109
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq Version 1.5
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/410,784A
/ FILING DATE: 24-MAR-1995
/ CLASSIFICATION: 800
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Heine, Ph.D., Holliday C
/ REGISTRATION NUMBER: 34,346
/ REFERENCE/DOCKET NUMBER: ISU-002XX
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617-542-2290
/ TELEFAX: 617-451-0313
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2523 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: Genomic DNA
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE:
/ ORIGINAL SOURCE:
/ US-08-410-784A-3

Query Match      4.2%; Score 58.6; DB 2; Length 2523;
Best Local Similarity 71.4%; Pred. No. 0.00013;
Matches 105; Conservative 0; Mismatches 39; Indels 3; Gaps 2;

QY 774 AAAATCTTCTGATTTTAAAGCTAGTTGGCAACCTGTTCTTTCAAGAAATT 833
DB 279 ACATATACATATACATATATATAGGCTAGTTGCAATCCATTTTATCAAAAAG--TTT 222
QY 834 TGATTTTCAAAAAAATTAGTTATTTCTCTTATTAATAATAGAAAACCTTGAAAA 893
DB 221 TACATTTTCAAAATTAATGTTATTTCTCTTGA-AAAATAGAAATTTCTCAGAAAA 163
QY 894 ATAGAGTTGCCAGACTAGCCCTAGAT 920
DB 162 ATAGAGTTTACAACACTAGCTTAAAT 136

RESULT 6
US-08-487-826B-13/C
/ Sequence 13, Application US/08487826B
/ Patent No. 5993827
/ GENERAL INFORMATION:
/ APPLICANT: Sim, Kim L.
/ APPLICANT: Chlitis, Chetan
/ APPLICANT: Miller, Louis H.
/ APPLICANT: Peterson, David S.
/ APPLICANT: Su, Xin-zhaun
/ APPLICANT: Wellens, Thomas E.
/ TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
/ TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
/ NUMBER OF SEQUENCES: 45
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Knobbe Martens Olson & Bear
/ STREET: 620 Newport Center Drive 16th Floor
/ CITY: Newport Beach
/ STATE: California
/ COUNTRY: US
```

ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israel, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121,001CPI
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 19124 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-487-826B-13

Query Match 4.1%; Score 57; DB 2; Length 19124;
Best Local Similarity 50.2%; Pred. No. 0.00073;
Matches 141; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY 614 TAAGCATGACGCTTCAGATTCTTTCTTCATCTCTGTATTTGTTATTTATAT 673
DB 15677 TAATGTTTTTTTTCTTCTTTCTTTGTTTAAATCAATTTTTTTTATAT 15618
QY 674 ACATTTCTCTCTCAATAGAGTGAATTTCCGATTTTAAATGACTATAAGT 723
DB 15617 AAAATTTTTTTTAAATTTTTTTTGATATCTTTTCATTTTATTTCAAAAATTTA 15558
QY 734 CATTTTAAATGAAGACGACGATGCTGATCTCGTTCAAAATCTTTCTGATTTTT 793
DB 15557 TATTTATATTAATTTTATTTTAAATTTTCCCTTTTTTTTTTTTTTTT 15498
QY 794 TAAGACTACTTGGCAACCTGTTCTTCAAGATTGATTTTCAAAAAAATTT 853
DB 15497 TTTTATTTAAATTTTTTTTTTATTTTCAATTTTCTTTTCAATTTTAAATTT 15438
QY 854 AGTTATTTCTCTTATAAATAGAAAACACTTAGAAAA 894
DB 15437 GTTTTATATTTCTTTTATTAATTAATGATATATATAA 15397

RESULT 7
US-09-313-294A-5397
Sequence 5397, Application US/09313294A
Patent No. 6476212
GENERAL INFORMATION:
APPLICANT: Laligudi, Raghunath V.
APPLICANT: Ito, Laura Y.
APPLICANT: Sherman, Bradley K.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
FILE REFERENCE: PL-0017 US
CURRENT APPLICATION NUMBER: US/09/313,294A
NUMBER OF SEQ ID NOS: 7600
SOFTWARE: PERL Program
SEQ ID NO 5397
LENGTH: 279
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc_feature

OTHER INFORMATION: Incyte ID No. 6476212 700350078H1
NAME/KEY: unsure
LOCATION: 10 12, 185-186, 204, 253, 274, 278
OTHER INFORMATION: a, c, g, or other
US-09-313-294A-5397

Query Match 4.0%; Score 55.8; DB 4; Length 279;
Best Local Similarity 69.0%; Pred. No. 0.00029;
Matches 89; Conservative 0; Mismatches 39; Indels 1; Gaps 1;

QY 789 TTTTAAAGCTAGTTGGCAACCTGTTCTTCAAGAAATTTGATTTTCAAAA 848
DB 3 TTCCCTAANGCCTAGTTGGAAACCCCATTTTCCACGCGTTTTCATTTCCAAAGG 62
QY 849 AATTAGTTATTTCTCTTTAT-AAATAGAAAACACTTAGAAAATAGATTGCA 907
DB 63 AAGTTAGAACATTTTCCCTTGGGAAAATAGAAATCTTTGGGAAAATGAGTTCCCAA 122
QY 908 CTAGCCCTA 916
DB 123 CTAGCCCTA 131

RESULT 8
US-09-902-540-1357/c
Sequence 1357, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:

APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 1357
LENGTH: 612
TYPE: DNA
ORGANISM: Myxococcus xanthus
FEATURE:
NAME/KEY: unsure
LOCATION: (1)-(612)
OTHER INFORMATION: unsure at all n locations
US-09-902-540-1357

Query Match 4.0%; Score 55.6; DB 4; Length 612;
Best Local Similarity 51.4%; Pred. No. 0.00043;
Matches 127; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY 626 TTCAGATTTCTTTCATCTGTTATTTGTTATTTGTTTAAATCAATTTCTCT 685
DB 560 TTTTATATTT 501
QY 686 CTACAAATGAGTATTTCTCCGATTTTAAATGACTATAAGTATTTTATATA 745
DB 500 TTTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTTAAATTTTTTTTA 441
QY 746 AAGACGAGATGCTGAGATCTCGTCAAAATCTTTGATTTTAAAGACTAGTT 805
DB 440 ATTATATTTATTT 381
QY 806 TGGCAACCTGTTCTTCAAGAAATTTGATTTTCAAAAAAATAGTTATTTCT 865
DB 380 TTTTATTTTTTTTTTTTTTTTATTTATTTTATTTTATTTTATTTTATTTT 321
QY 866 CTTATA 872
DB 320 TATTTTA 314

RESULT 9
US-08-880-499-1/c
; Sequence 1, Application US/08880499
; Patent No. 6037523
; GENERAL INFORMATION:
; APPLICANT: Albertson, Marc C.
; APPLICANT: Fox, Tim W.
; APPLICANT: Carl, Garnaat W.
; APPLICANT: Huffman, Gary A.
; APPLICANT: Kendall, Timmy L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION
; TITLE OF INVENTION: AND METHOD OF USING SAME
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
; STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.
; STREET: Box 1000
; CITY: Johnston
; STATE: Iowa
; COUNTRY: USA
; ZIP: 50131
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/880,499
; FILING DATE: CONCURRENTLY HERewith
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Sweeney, Patricia A.
; REGISTRATION NUMBER: 32,733
; REFERENCE/DOCKET NUMBER: 0578
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (515) 248-4800
; TELEFAX: (515) 248-4844
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1394 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-880-499-1
Query Match 4.0%; Score 55.6; DB 3; Length 1394;
Best Local Similarity 55.8%; Pred. No. 0.00059;
Matches 106; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
QY 761 TAGATTCTGCTCAAAAATCTTTCTGATTTTAAAGAGCTAGTTGGCAACCCGTTTC 820
DB 950 TACACAGTATGTATTTATTTGGAACAATCTAGGGCTAGCTGGCAACTCTATTTT 891
QY 821 TTTCAGAAATTTGATTTTCAAAAAAATAGTTATTTCTCTTTATAAATAGAA 880
DB 890 TCTAAGTGTCTTATTTTATAAGAGAAATTAACCTAATTTTGTGAAAAATCAAAA 831
QY 881 AACACTTAGAAAAATAGAGTTGCCAGACTAGCCCTAGAAATGTTTCCCAATTAATACAA 940
DB 830 TTCTTTGAAAGAAAGAGGGTTCACAACTAGCTCTTAATAAATCAGAAAGATTTTGAA 771
QY 941 TCACGTGTGA 950
DB 770 CGAGATCTA 761

RESULT 10
US-08-880-499-2/c
; Sequence 2, Application US/08880499
; Patent No. 6037523
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; APPLICANT: Albertson, Marc C.
; APPLICANT: Fox, Tim W.
; APPLICANT: Carl, Garnaat W.
; APPLICANT: Huffman, Gary A.
; APPLICANT: Kendall, Timmy L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION
; TITLE OF INVENTION: AND METHOD OF USING SAME
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
; STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.
; STREET: Box 1000
; CITY: Johnston
; STATE: Iowa
; COUNTRY: USA
; ZIP: 50131
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/880,499
; FILING DATE: CONCURRENTLY HERewith
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Sweeney, Patricia A.
; REGISTRATION NUMBER: 32,733
; REFERENCE/DOCKET NUMBER: 0578
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (515) 248-4800
; TELEFAX: (515) 248-4844
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1394 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-880-499-2
Query Match 4.0%; Score 55.6; DB 3; Length 1394;
Best Local Similarity 55.8%; Pred. No. 0.00059;
Matches 106; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
QY 761 TAGATTCTGCTCAAAAATCTTTCTGATTTTAAAGAGCTAGTTGGCAACCCGTTTC 820
DB 950 TACACAGTATGTATTTATTTGGAACAATCTAGGGCTAGCTGGCAACTCTATTTT 891
QY 821 TTTCAGAAATTTGATTTTCAAAAAAATAGTTATTTCTCTTTATAAATAGAA 880
DB 890 TCTAAGTGTCTTATTTTATAAGAGAAATTAACCTAATTTTGTGAAAAATCAAAA 831
QY 881 AACACTTAGAAAAATAGAGTTGCCAGACTAGCCCTAGAAATGTTTCCCAATTAATACAA 940
DB 830 TTCTTTGAAAGAAAGAGGGTTCACAACTAGCTCTTAATAAATCAGAAAGATTTTGAA 771
QY 941 TCACGTGTGA 950
DB 770 CGAGATCTA 761

RESULT 11
US-09-806-708B-23
; Sequence 23, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133

; PRIOR FILING DATE: 1999-08-04
 ; NUMBER OF SEQ ID NOS: 23
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 23
 ; LENGTH: 1055
 ; TYPE: DNA
 ; ORGANISM: Artificial sequence
 ; FEATURE:
 ; NAME/KEY: promoter
 ; LOCATION: (1)..(1055)
 ; OTHER INFORMATION: consensus sequence of A.T. and L.A. FAE1 promoters
 US-09-806-708B-23

Query Match 4.0%; Score 55.4; DB 4; Length 1055;
 Best Local Similarity 22.5%; Pred. No. 0.0006;
 Matches 182; Conservative 181; Mismatches 430; Indels 15; Gaps 3;

QY 292 TCACCTATATTATCTTGAATATTATTTGAAATAACAACTTATAC 351
 DB 37 YCANNTGKRCYARWGMHTTAYTTATKMGTTGAWMTWMAWKTRKMCSTAMNNAW 96
 QY 352 TTTTGTGAGGGCCCTCAG-CATAGATTTTCCTTAGGGCCAGAAATCGAGACCGCC 410
 DB 97 TTWCTARKMTGTRWMTKTNNAATGTRWMTGYMTNNNGSTWMTMARVYTRRMWCYTAM 156
 QY 411 ATGCTTAGTGTCCACTATTGGCACTACCCAGAACAGATTAAATAAATAACCAAGTAA 470
 DB 157 WYASVAGNSTRRTTYTWMKMKCKKKSARABATRRARVMDAWYAMARRGTWKAMA 216
 QY 471 CTATCCACTCGAAGCTATCATGTATGT--TTAAGAAACATCATTAACCAACCGA 527
 DB 217 YAAWMTNNNNNAKAKCKRATTWGMRKSNCTCTTAGCTTTCATCCWATTCAGWMTKK 276
 QY 528 TCCTCTTAAAAAAACAACATATTTCGAAAGAGACAAATTGTACAGTTTACAAACATC 587
 DB 277 WKTMKTSAAAGMTWMTNNNNNTTTTKAAMVAAABMMWMSATTTWMAAAWTSRTWYGR 336
 QY 588 TAAAGGACGACAAATATATGAAAGTAAGCTATGACGTTACAGTTTCTTTTCATTC 647
 DB 337 KTMNNNNNGTTCWTRWAMTWKMKMKGTNNNNNGRTYGTWTKKVAATTTAAKANN 396
 QY 648 TTGTTATTTGTTATGTTTATATATATATCTTCTCTTACATAGATGATTTCTT 707
 DB 397 TTAAMWKCTTNNNTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 456
 QY 708 CCGATTTTAAATAAGCTATA-----AAGCATTTTATATTAAGACAGCAT 756
 DB 457 NNTKTMKACTWTTTBRCTTANNNTAATWTKSSANCTSRTRKTKMCRAGSKTAAWGRA 516
 QY 757 GTCTAGATTCGTTCAAAAATCTTTCATTTTAAAGCTAGTTGGCAACCTG 816
 DB 517 YABAATGKMTAMVMTWMTWYVYAGAAWMTAMWMTSATCYCAPAATTAAGCAGAGS 576
 QY 817 TTTCTTTCAAGAATTTGATTTTTCACAAAAAATTAAGTTATTTCTCTTATAAAT 876
 DB 577 TAKGNNNNNNNCCATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 636
 QY 877 AGAAAACTTAGAAAAATAGAGTGCAGACTAGCCCTAGAAATGTTTCCCAATAAT 936
 DB 637 NATKTAATWTTNNNNNAGTWTNNNNNNAAGAAATTAATTAATTAATTAATTAAT 696
 QY 937 ACATCACTGTGATAATTAATTTGGCCAGCCCATTAATTAATTAATTAATTAATTAAT 996
 DB 697 AYRAAAYTTRTANNGCTTTTNNNTTGGMRNTTAARGMANNNNNNNNNNNNGACVA 756
 QY 997 CGAGCGAAACCAATCGAGCTATTTCTAGATTGTAATAAGGAGAGAGAGAGAG 1056
 DB 757 WRTTATACGTTNNNNNNNNNNNAATTTTATTTTWTTRKANNNNNNNNNAAYYGA 816
 QY 1057 AATCAGTTTAAATGCTATGCTCCTGAG 1084
 DB 817 KNTTWCCTCKAMWKAMATGAATTTNAG 844

RESULT 12
 US-09-004-056-1/c
 ; Sequence 1, Application US/09004056A
 ; Patent No. 656586
 ; GENERAL INFORMATION:
 ; APPLICANT: Calgene LLC
 ; TITLE OF INVENTION: Plant Expansin Promoter Sequences
 ; FILE REFERENCE: 125
 ; CURRENT APPLICATION NUMBER: US/09/004, 056A
 ; CURRENT FILING DATE: 1998-01-07
 ; EARLIER APPLICATION NUMBER: 60034914
 ; EARLIER FILING DATE: 1997-07-01
 ; NUMBER OF SEQ ID NOS: 1
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 2614
 ; TYPE: DNA
 ; ORGANISM: Geesypium hirsutum
 ; FEATURE:
 ; NAME/KEY: promoter
 ; LOCATION: (930)
 ; OTHER INFORMATION: unknown nucleotide
 ; FEATURE:
 ; NAME/KEY: promoter
 ; LOCATION: (947)
 ; OTHER INFORMATION: unknown nucleotide
 ; FEATURE:
 ; NAME/KEY: promoter
 ; LOCATION: (956)
 ; OTHER INFORMATION: unknown nucleotide
 US-09-004-056-1

Query Match 4.0%; Score 55.2; DB 4; Length 2614;
 Best Local Similarity 54.2%; Pred. No. 0.00095;
 Matches 155; Conservative 0; Mismatches 128; Indels 3; Gaps 2;

QY 611 AGTAAGCTATGACGCTGATTTTCTTTTCATCTTGTATTTGTTATGTTTAA 670
 DB 447 AGTTAGTTGTTATCAATTTTTCATATTAATTTAGTTTATTTATTTCTAAATTA 388
 QY 671 TATACATTTCTCTTACATATAGATTTTCTCGATTTTAAATGACATATA 730
 DB 387 TGTGCAAAATGAACCTTTATTTATATTTTAAATATATTTGATTAAT-TTTAAAG 329
 QY 731 AGCATTTTATATAGACAGCATGCTGATTCGTTCAAAAATCTTCTGATTT 790
 DB 328 TATTTTCAATATATATTTTCAGGAAACATTAATTTGCAATACGAATTTTGAGATT 269
 QY 791 TTTTAAGACTAGTTGGCAACCTGTTCTTTCAAAAGATTTGATTTTTCAAAAA 850
 DB 268 TTTAAATGTATGATTTTAAATTTCTTATTAATTAATTAATTAATTAATTAAT 211
 QY 851 ATTAGTTATTTCTCTTATTAATTAAGAAACACTTAAGAAAAATA 896
 DB 210 ATTAGTTAATGAAATTTTAAATTTTAAATTAATTAAGAAAAATA 165

RESULT 13
 US-09-640-173-53
 ; Sequence 53, Application US/09640173
 ; Patent No. 6613515
 ; GENERAL INFORMATION:
 ; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Stolk, John A.
 ; TITLE OF INVENTION: OVARIAN TUMOR SEQUENCES AND
 ; TITLE OF INVENTION: METHODS OF USE THEREFOR
 ; FILE REFERENCE: 210121.484C2
 ; CURRENT APPLICATION NUMBER: US/09/640,173
 ; CURRENT FILING DATE: 2000-08-15
 ; NUMBER OF SEQ ID NOS: 196
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 53

Search completed: September 15, 2005, 08:25:12
Job time : 212.859 secs

This Page Blank (uspro)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 14, 2005, 19:14:19 ; Search time 23.7481 Seconds
(without alignments)
10817.505 Million cell updates/sec

Title: US-10-713-381-2_COPY_1155_1311

Perfect score: 157
1 cgcgcacatccacatggcat.....cgcacatcgttcccggaac 157

Sequence: 157

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B COMB.seq: *
3: /cgn2_6/ptodata/1/ina/5A COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PTUS COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	157	100.0	1394	3 US-08-880-499-1	Sequence 1, Appl1
2	157	100.0	1394	3 US-08-880-499-2	Sequence 2, Appl1
3	30.4	19.4	385136	4 US-09-949-016-16073	Sequence 16073, A
4	29.6	18.9	1185	4 US-09-543-681A-1081	Sequence 1081, Ap
5	29.2	18.6	1830121	4 US-09-557-884-1	Sequence 1, Appl1
6	29.2	18.3	64190	4 US-09-643-990A-1	Sequence 14712, A
7	28.8	18.3	64190	4 US-09-949-016-14712	Sequence 14712, A
8	28.8	17.8	11032	4 US-09-949-016-15502	Sequence 15502, A
9	28	17.8	41815	4 US-09-949-016-17447	Sequence 17447, A
10	28	17.7	601	4 US-09-949-016-191201	Sequence 191201, A
11	27.8	17.7	11490	4 US-09-949-016-16149	Sequence 16149, A
12	27.8	17.7	14952	4 US-09-949-016-16740	Sequence 16740, A
13	27.4	17.5	1413	4 US-09-710-279-1813	Sequence 1813, Ap
14	27.4	17.5	1438	4 US-09-134-001C-1413	Sequence 1413, Ap
15	27.4	17.5	1794	4 US-09-270-767-10489	Sequence 10489, A
16	27.4	17.5	2472	4 US-08-425-295A-2	Sequence 2, Appl1
17	27.4	17.5	2926	4 US-09-710-279-3747	Sequence 3747, Ap
18	27.4	17.5	3014	4 US-09-710-279-3680	Sequence 3680, Ap
19	27.4	17.5	3539	4 US-09-949-016-2556	Sequence 2556, Ap
20	27.4	17.3	3445	4 US-09-949-016-2556	Sequence 2556, Ap
21	27.2	17.3	3588	4 US-09-566-821-23	Sequence 23, Appl1
22	27.2	17.3	6751	1 US-07-882-925A-5	Sequence 5, Appl1
23	27.2	17.3	6751	1 US-08-184-012C-5	Sequence 12703, A
24	27.2	17.3	8939	4 US-09-949-016-12703	Sequence 12703, A
25	27.2	17.3	9076	4 US-09-949-016-17368	Sequence 17368, A
26	27.2	17.3	9076	4 US-09-949-016-17368	Sequence 17368, A
27	27.2	17.3	48934	4 US-09-949-016-14091	Sequence 14091, A

28	27.2	17.3	60593	4 US-09-949-016-13779	Sequence 13779, A
29	27.2	17.3	62776	4 US-09-949-016-17576	Sequence 17576, A
30	27.2	17.3	100848	4 US-09-596-002-39	Sequence 39, Appl1
31	27.2	17.3	158735	4 US-09-949-016-11989	Sequence 11989, A
32	27.2	17.3	158735	4 US-09-949-016-17130	Sequence 17130, A
33	27.2	17.2	879	3 US-09-134-001C-678	Sequence 678, Appl
34	27	17.2	44180	4 US-09-949-016-12988	Sequence 12088, A
35	27	17.2	47030	4 US-09-949-016-13037	Sequence 13037, A
36	27	17.2	47030	4 US-09-949-016-13038	Sequence 13038, A
37	27	17.2	47030	4 US-09-949-016-15039	Sequence 15039, A
38	27	17.2	47030	4 US-09-949-016-15040	Sequence 15040, A
39	27	17.2	194537	4 US-09-949-016-12928	Sequence 12928, A
40	27	17.2	201529	4 US-09-949-016-12740	Sequence 12740, A
41	26.8	17.1	265	4 US-09-513-999C-8379	Sequence 8379, Ap
42	26.8	17.1	271	3 US-08-946-026-9	Sequence 9, Appl1
43	26.8	17.1	3114	3 US-08-946-026-12	Sequence 12, Appl1
44	26.8	17.1	156942	4 US-09-949-016-12227	Sequence 12227, A
45	26.8	17.1	156950	4 US-09-949-016-15946	Sequence 15946, A

ALIGNMENTS

RESULT 1
US-08-880-499-1
Sequence 1, Application US/08880499

Patent No. 6037523

GENERAL INFORMATION:

APPLICANT: Albertson, Marc C.

APPLICANT: Fox, Tim W.

APPLICANT: Carl, Garmaat W.

APPLICANT: Hufman, Timmy L.

APPLICANT: Kendall, Gary L.

TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION

TITLE OF INVENTION: AND METHOD OF USING SAME

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.

STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.

STREET: Box 1000

CITY: Johnston

STATE: Iowa

COUNTRY: USA

ZIP: 50131

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/880,499

FILING DATE: CONCURRENTLY HEREWITH

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Sweeney, Patricia A.

REGISTRATION NUMBER: 32,733

REFERENCE/DOCKET NUMBER: 0578

TELECOMMUNICATION INFORMATION:

TELEPHONE: (515) 248-4800

TELEFAX: (515) 248-4844

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1394 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-880-499-1

Query Match 100.0%; Score 157; DB 3; Length 1394;
Best Local Similarity 100.0%; Pred. No. 5.6e+45;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGATCTACATGAGCATACATGCTGTGTCACCGTTCGTGTTCCATGCTCC 60
DB 1155 CGTGATCTACATGAGCATACATGCTGTGTCACCGTTCGTGTTCCATGCTCC 1214
QY 61 AAGCTTGGCTTATTCGAACCAAGAGATACCTACTCCAAACATTCATCTTACTG 120
DB 1215 AAGCTTGGCTTATTCGAACCAAGAGATACCTACTCCAAACATTCATCTTACTG 1274
QY 121 CAACCTTCATGCAACGACATATGTTCTCGAAC 157
DB 1275 CAACCTTCATGCAACGACATATGTTCTCGAAC 1311

RESULT 2
US-08-880-499-2
Sequence 2, Application US/08880499
Patent No. 6037523
GENERAL INFORMATION:
APPLICANT: Alberson, Marc C.
APPLICANT: Fox, Tim W.
APPLICANT: Carl, Garraat W.
APPLICANT: Huffman, Gary A.
APPLICANT: Kendall, Timmy L.
TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION
TITLE OF INVENTION: AND METHOD OF USING SAME
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESSES:
ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.
CITY: Johnston
STATE: Iowa
COUNTRY: USA
ZIP: 50131
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/880,499
FILING DATE: CONCURRENTLY HEREWITH
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Sweeney, Patricia A.
REGISTRATION NUMBER: 32,733
REFERENCE/DOCKET NUMBER: 0578
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 248-4800
TELEFAX: (515) 248-4844
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1394 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-880-499-2

Query Match 100.0%; Score 157; DB 3; Length 1394;
Best Local Similarity 100.0%; Pred. No. 5, 6e-45;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGATCTACATGAGCATACATGCTGTGTCACCGTTCGTGTTCCATGCTCC 60
DB 1155 CGTGATCTACATGAGCATACATGCTGTGTCACCGTTCGTGTTCCATGCTCC 1214
QY 61 AAGCTTGGCTTATTCGAACCAAGAGATACCTACTCCAAACATTCATCTTACTG 120
DB 1215 AAGCTTGGCTTATTCGAACCAAGAGATACCTACTCCAAACATTCATCTTACTG 1274
QY 121 CAACCTTCATGCAACGACATATGTTCTCGAAC 157
DB 1275 CAACCTTCATGCAACGACATATGTTCTCGAAC 1311

DB 1275 CAACCTTCATGCAACGACATATGTTCTCGAAC 1311

RESULT 3
US-09-949-016-16073/C
Sequence 16073, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 16073
LENGTH: 385136
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(385136)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16073

Query Match 19.4%; Score 30.4; DB 4; Length 385136;
Best Local Similarity 57.3%; Pred. No. 14;
Matches 55; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 55 TCGTCCAGCGCTTGGCTTATTCGAACCAAGAGATACCTACTCCAAACATTCATCTTA 114
DB 253908 TCGTCCAGCGCTTGGCTTATTCGAACCAAGAGATACCTACTCCAAACATTCATCTTA 253849
QY 115 CTCATGCAACTTCCTGCAACGACGACATATGTTT 150
DB 253848 TCCATTTATCTCCGACACACCATATTTT 253813

RESULT 4
US-09-543-681A-1081/C
Sequence 1081, Application US/09543681A
Patent No. 6605709
GENERAL INFORMATION:
APPLICANT: GARY BRETON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
FILE REFERENCE: 2709,1002-001
CURRENT APPLICATION NUMBER: US/09/543,681A
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: 60/128,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO 1081
LENGTH: 1185
TYPE: DNA
ORGANISM: Proteus mirabilis
US-09-543-681A-1081

Query Match 18.9%; Score 29.6; DB 4; Length 1185;
Best Local Similarity 59.5%; Pred. No. 2, 4;
Matches 50; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 35 CAACGCTTGTGTTGTCATTCGTCAGAGCTTGTCTATTCGAACCAAGAGATACCTTA 94
DB 940 CAACGATTTGTTTGTGTCACCAAGAAAGACCTTAACCATTTTCTCCAAAGACATATCTTA 881

```

QY      95 CTCCCAACATCCATCTTACTCA 118
          |||| | | | | | |
Db      880 ATCCCTAGATGCTTTTATATA 857

```

RESULT 5
115-09-55

US-09-557-884-1/C
Sequence 1, Application US/09557884
Patent No. 6505581
GENERAL INFORMATION:
APPLICANT: Pleischmann et al.
TITLE OF INVENTION: The Nucleotide sequence of the Haemophilus influenzae Rd Genome, Fragments Thereof, and Uses Thereof
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS v6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/557,884
FILING DATE: 25-Apr-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/476,102
FILING DATE: JUN-5-1995
ATTORNEY/AGENT INFORMATION:
NAME: Michelle S. Marks
REGISTRATION NUMBER: 41,971
REFERENCE/DOCKET NUMBER: PB186P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1830121 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1

RESULT 6
US-09-643-990A-1/c
; Sequence 1, Application US/09643990A
; Patent No. 6528389
; GENERAL INFORMATION:
; APPLICANT: Robert D. Fleischmann

```

1 Mark D. Adams
2 Owen White
3 Hamilton O. Smith
4 J. Craig Venter
5
6 TITLE OF INVENTION: The Nucleotide sequence of
7 the Haemophilus influenzae Rd Genome, Fragments
8 Thereof, and Uses Thereof
9
10 NUMBER OF SEQUENCES: 1
11
12 CORRESPONDENCE ADDRESS:
13 ADDRESSEE: Human Genome Sciences, Inc.
14 STREET: 9410 Key West Avenue
15 CITY: Rockville,
16 STATE: MD
17 COUNTRY: USA
18 ZIP: 20850
19
20 COMPUTER READABLE FORM:
21 MEDIUM TYPE: 3 1/2 inch diskette
22
23 COMPUTER: Dell Pentium
24 OPERATING SYSTEM: MS DOS v6.22
25 SOFTWARE: ASCII Text
26
27 CURRENT APPLICATION DATA:
28 APPLICATION NUMBER: US/09/643, 990A
29 FILING DATE: 23-Aug-2000
30 CLASSIFICATION: <Unknown>
31
32 PRIOR APPLICATION DATA:
33 APPLICATION NUMBER: 08/487, 429
34 FILING DATE: 1995-06-07
35 APPLICATION NUMBER: 08/426, 787
36 FILING DATE: 1995-04-21
37
38 ATTORNEY/AGENT INFORMATION:
39 NAME: Kenley K. Hoover
40 REGISTRATION NUMBER: 40,302
41 REFERENCE/DOCKET NUMBER: P186P1C1
42 TELECOMMUNICATION INFORMATION:
43 TELEPHONE: 301-610-5790
44 TELEFAX: 310-509-8439
45
46 INFORMATION FOR SEQ ID NO: 1:
47
48 SEQUENCE CHARACTERISTICS:
49 LENGTH: 183021 base pairs
50 TYPE: nucleic acid
51 STRANDEDNESS: double
52 TOPOLOGY: linear
53
54 SEQUENCE DESCRIPTION: SEQ ID NO: 1:
55
56 US-09-643-990A-1
57
58 Query Match 18.6%; Score 29.2; DB 4; Length 183021;
59 Best Local Similarity 51.5%; Pred. No. 66;
60 Matches 67; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

```

RESULT 7
 US-09-949-016-14712/c
 : Sequence 14712, Application US/09949016
 : Patent No. 6812319
 : GENERAL INFORMATION:
 : APPLICANT: VENTER, J. Craig et al.
 : TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 : WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 : FILE REFERENCE: CL001037
 : CURRENT APPLICATION NUMBER: US/09/949,016
 : CURRENT FILING DATE: 2000-04-14

```

; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14712
; LENGTH: 64190
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14712

Query Match
Best Local Similarity 18.3%; Score 28.8; DB 4; Length 64190;
Pred. No. 24;
Matches 45; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 55 TCCTCCAGCCTTGCCTATTCGACCAAGAGATACCTTCCCAACATCCATCTTA 114
DB 355 TCATCCAAATGTTCCACTACTAATCCATGATGGAAGTACTAATCAAGCAATATT 296
QY 115 CTCATGCAACTT 126
DB 295 CTCATGTAACAT 284

RESULT 8
US-09-949-016-14713/c
; Sequence 14713, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14713
; LENGTH: 64190
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14713

Query Match
Best Local Similarity 18.3%; Score 28.8; DB 4; Length 64190;
Pred. No. 24;
Matches 45; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 55 TCCTCCAGCCTTGCCTATTCGACCAAGAGATACCTTCCCAACATCCATCTTA 114
DB 355 TCATCCAAATGTTCCACTACTAATCCATGATGGAAGTACTAATCAAGCAATATT 296
QY 115 CTCATGCAACTT 126
DB 295 CTCATGTAACAT 284

RESULT 9
US-09-949-016-15502
; Sequence 15502, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15502
; LENGTH: 11022
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15502

Query Match
Best Local Similarity 17.8%; Score 28; DB 4; Length 11022;
Pred. No. 22;
Matches 43; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 15 TGGCATCTACATGCTTGTTCACCGCTTCGCTTGTTCATGTCACAGCTTGCTATT 74
DB 4432 TGGCTTATTTCATGCTGTGTTTAAAGGTTAATGTTGATGATCATTCATTAGCTTGCAAG 4491
QY 75 CTGAACCA 82
DB 4492 CTGAATCA 4499

RESULT 10
US-09-949-016-17447
; Sequence 17447, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17447
; LENGTH: 41815
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17447

Query Match
Best Local Similarity 17.8%; Score 28; DB 4; Length 41815;
Pred. No. 38;
Matches 49; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 62 AGCCTTGCTATTTCGACCAAGAGATACCTTCCCAACATCTTACTCATGTC 121
DB 33056 AGACATGGAGATTACACACATCCGATAGCTCTCTAAGCATCCCTGCTCATTC 33117
QY 122 AACTTCATGCAACACGCACATA 145
DB 33118 ATGTCATAGAGACATGATGTA 33141

RESULT 11
US-09-949-016-191201/c
; Sequence 191201, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 191201
; LENGTH: 11022
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-191201/c
```

```

; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15502
; LENGTH: 11022
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15502

Query Match
Best Local Similarity 17.8%; Score 28; DB 4; Length 11022;
Pred. No. 22;
Matches 43; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 15 TGGCATCTACATGCTTGTTCACCGCTTCGCTTGTTCATGTCACAGCTTGCTATT 74
DB 4432 TGGCTTATTTCATGCTGTGTTTAAAGGTTAATGTTGATGATCATTCATTAGCTTGCAAG 4491
QY 75 CTGAACCA 82
DB 4492 CTGAATCA 4499

RESULT 10
US-09-949-016-17447
; Sequence 17447, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17447
; LENGTH: 41815
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17447

Query Match
Best Local Similarity 17.8%; Score 28; DB 4; Length 41815;
Pred. No. 38;
Matches 49; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 62 AGCCTTGCTATTTCGACCAAGAGATACCTTCCCAACATCTTACTCATGTC 121
DB 33056 AGACATGGAGATTACACACATCCGATAGCTCTCTAAGCATCCCTGCTCATTC 33117
QY 122 AACTTCATGCAACACGCACATA 145
DB 33118 ATGTCATAGAGACATGATGTA 33141

RESULT 11
US-09-949-016-191201/c
; Sequence 191201, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 191201
; LENGTH: 11022
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-191201/c
```

```
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 191201
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-191201
```

Query Match 17.7%; Score 27.8; DB 4; Length 601;
Best Local Similarity 58.0%; Pred. No. 7.9;
Matches 47; Conservative 1; Mismatches 33; Indels 0; Gaps 0;

```
QY 57 GTCCAGCCTTGCTATTCTGAACAGAGATACCTACCCCAACATCTTACT 116
DB 322 GTCTGACCTTGATGACACRTGACGCGCCCAACACACACACTCTCTCAT 263
QY 117 CATGCACTTCATGCAACA 137
DB 262 CATGCAATATGATGAAATA 242
```

```
RESULT 12
US-09-949-016-14149/c
Sequence 14149, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14149
LENGTH: 11490
TYPE: DNA
ORGANISM: Human
US-09-949-016-14149
```

Query Match 17.7%; Score 27.8; DB 4; Length 11490;
Best Local Similarity 52.1%; Pred. No. 26;
Matches 62; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

```
QY 38 CCGTGTCTTGTTCATCTGTCAGCCTTGCTATTGGAACCAAGAGATCTACTC 97
DB 5642 CAGGGCTCTGCTTCCCTGCCCCCTGCCCCCACCCTCACAGAGAGAGCCACCC 5583
QY 98 CCAACATCATCTTACTGATGCACTTCATGCAACGACGACATATGTTCCGAA 156
DB 5582 CTTGACACCTGATGACATCCGCTTAGTGCGAGGCGAGTGAAGATGTTGCCAAA 5524
```

RESULT 13
US-09-949-016-16740

```
Sequence 16740, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 16740
LENGTH: 14952
TYPE: DNA
ORGANISM: Human
US-09-949-016-16740
```

Query Match 17.7%; Score 27.8; DB 4; Length 14952;
Best Local Similarity 59.5%; Pred. No. 29;
Matches 47; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

```
QY 67 TGCCTATTCTGAACAGAGATACCTACTCCCAACATCTTACTGCAACTT 126
DB 70 TGTCAACCGTGAATCCAGAGTAAATGTTTGTGAAATTCCTCTCAGGTAACTA 129
QY 127 CCATGCAACAGCCACATA 145
DB 130 GGTTTTACACACACACA 148
```

```
RESULT 14
US-09-710-279-1813/c
Sequence 1813, Application US/09710279
Patent No. 6703492
GENERAL INFORMATION:
APPLICANT: KIMBERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: PUS480US
CURRENT APPLICATION NUMBER: US/09/710,279
CURRENT FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: 60/164,258
PRIOR FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1813
LENGTH: 1413
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-710-279-1813
```

Query Match 17.5%; Score 27.4; DB 4; Length 1413;
Best Local Similarity 55.9%; Pred. No. 15;
Matches 52; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

```
QY 14 ATGGATACTACATCTGTTCAACCGTTCGCTTGTTCATGTCGCAAGCCTTGCTAT 73
DB 1213 ATTGAATCTACGTCTGTTCCACAGATTGTTATCTTCATCTGATATATCATCATA 1154
QY 74 TCTGAACCAAGAGATACCTACTCCCAACAT 106
DB 1153 CAAGAATGCAATATATCTGTACTCTCTAT 1121
```

RESULT 15

US-09-134-001C-1413/C
; Sequence 1413 Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1413
; LENGTH: 1428
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1413

Query Match 17.5%; Score 27.4; DB 3; Length 1428;
Best Local Similarity 55.9%; Pred. No. 15;
Matches 52; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY	14	ATGCGATPACTACATGCTGTGTTCAACGCTGCTGTGTTCCATGCTCCAGCCTTGCTTAT	73
DB	1228	ATTGAATTTCTACGCTCGCTCGTTCACAGTTTGTTATCTTCATCTGATTAATTCATCCATAC	1169
QY	74	TCTGAACCAAGAGATACCTACTCCCAACAAT	106
DB	1168	CAAGAAATCGCAATATATCTTGTAACTCTCTAT	1136

Search completed: September 15, 2005, 08:25:20
Job time : 28.7481 secs


```
Db 1 CCAATGCTCTATGAAAAAGATGAGTCAATGTCCTATATCCGTTTCTTAGAGTCC 60
Qy 61 CTTCTCTGCTTATTAAGTCACTGAATCGGGGTTTACAAAAAATTCCAGCGGTCATGAT 120
Db 61 CTTCTCTGCTTATTAAGTCACTGAATCGGGGTTTACAAAAAATTCCAGCGGTCATGAT 120
Qy 121 CTCGATGTTCCACTTCCTCCACCTCGGTCGATCATTTCTTGAGATCGGTCGATCCAT 180
Db 121 CTCGATGTTCCACTTCCTCCACCTCGGTCGATCATTTCTTGAGATCGGTCGATCCAT 180
Qy 181 CTGACCGAGGCCCATGACACCTTTGCGAGACCCCATCAAGGCGCTTTCGATGCGCA 240
Db 181 CTGACCGAGGCCCATGACACCTTTGCGAGACCCCATCAAGGCGCTTTCGATGCGCA 240
Qy 241 CGAGAGTATCGGGTCGTGTGATCCAGGGGATATATGTCCTCCCAATGCTCACTTAA 300
Db 241 CGAGAGTATCGGGTCGTGTGATCCAGGGGATATATGTCCTCCCAATGCTCACTTAA 300
Qy 301 TTATATTTCTTATGATATTTATTTTGAATAAATAAACAATTATACCTTTTGTGA 360
Db 301 TTATATTTCTTATGATATTTATTTTGAATAAATAAACAATTATACCTTTTGTGA 360
Qy 361 GGGCTTCAGATATGATTTTCGCTTAGGGCCAGAAAATGCGAGGACCAAGCTCTAGTG 420
Db 361 GGGCTTCAGATATGATTTTCGCTTAGGGCCAGAAAATGCGAGGACCAAGCTCTAGTG 420
Qy 421 TCCACTATTTGGCTCCAGGACAGATTTTAAATAAATAAACAAGTAACTTAACTACT 480
Db 421 TCCACTATTTGGCTCCAGGACAGATTTTAAATAAATAAACAAGTAACTTAACTACT 480
Qy 481 CGAAGCTATCATGATGTTTAAAGAAACATCTATTTAAACACGATCCTCTTAAATA 540
Db 481 CGAAGCTATCATGATGTTTAAAGAAACATCTATTTAAACACGATCCTCTTAAATA 540
Qy 541 CAAAGCATTTTGAAGAGACAAATATGTTACAGTTTACAAACATCTAAGACGACAA 600
Db 541 CAAAGCATTTTGAAGAGACAAATATGTTACAGTTTACAAACATCTAAGACGACAA 600
Qy 601 TTAATATGAAAGGTAAGCTATGAGTTCAGATTTTCTTTTCATCTGTTATTTTGT 660
Db 601 TTAATATGAAAGGTAAGCTATGAGTTCAGATTTTCTTTTCATCTGTTATTTTGT 660
Qy 661 ATTTGTTTATATACATTTTCTTCTCAATAGAGATTTTCTTCGATTTTATATA 720
Db 661 ATTTGTTTATATACATTTTCTTCTCAATAGAGATTTTCTTCGATTTTATATA 720
Qy 721 ATGACTATTAAGTCAATTTTATTAAGACGACGATGCTGATGATCTGTTCAAAAATC 780
Db 721 ATGACTATTAAGTCAATTTTATTAAGACGACGATGCTGATGATCTGTTCAAAAATC 780
Qy 781 TTTCTGATTTTAAAGAGTGTGGCAACCTGTTTCTTCAAGAAATTTGATTT 840
Db 781 TTTCTGATTTTAAAGAGTGTGGCAACCTGTTTCTTCAAGAAATTTGATTT 840
Qy 841 TTTCAAAAAAATTAAGTTATTTTCTTATTAATAATAGAAAACATTAGAAAAATAGAT 900
Db 841 TTTCAAAAAAATTAAGTTATTTTCTTATTAATAATAGAAAACATTAGAAAAATAGAT 900
Qy 901 TGGCAGACTAGCCCTTGAATGTTTTCCCAATAATTAATCAATCATCTGTATTAATTTG 960
Db 901 TGGCAGACTAGCCCTTGAATGTTTTCCCAATAATTAATCAATCATCTGTATTAATTTG 960
Qy 961 GCGAGCCCATTAATTAATTTAAACGAACTGAAATCGAGCGAAACCAATCTGAGCTAT 1020
Db 961 GCGAGCCCATTAATTAATTTAAACGAACTGAAATCGAGCGAAACCAATCTGAGCTAT 1020
Qy 1021 TTTCTAGATTAAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Db 1021 TTTCTAGATTAAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Qy 1081 TGAAGATGTCGGTTTGGCAACGATAGCCACCGTATCATAGCTCATAGGTGCTTAGCTCA 1140
Db 1081 TGAAGATGTCGGTTTGGCAACGATAGCCACCGTATCATAGCTCATAGGTGCTTAGCTCA 1140
```

```
Db 1081 TGAAGATGTCGGTTTGGCAACGATAGCCACCGTATCATAGCTCATAGGTGCTTAGCTCA 1140
Qy 1141 GGTTCGAGAGCTCTGTGTCATCTCAATGAGATCTCAATGCTTTTCAACGTTCCGTC 1200
Db 1141 GGTTCGAGAGCTCTGTGTCATCTCAATGAGATCTCAATGCTTTTCAACGTTCCGTC 1200
Qy 1201 TTGTTTCATTCGTCGAAGCTTTCCTATTTCTGAACCAAGAGATTAATCTCCAAACAAT 1260
Db 1201 TTGTTTCATTCGTCGAAGCTTTCCTATTTCTGAACCAAGAGATTAATCTCCAAACAAT 1260
Qy 1261 CCATTTTACTATGCACTTCCATGCAACACGACATATGTTTCTGTAACAGATCTATT 1320
Db 1261 CCATTTTACTATGCACTTCCATGCAACACGACATATGTTTCTGTAACAGATCTATT 1320
Qy 1321 AAAGATCAACAAGCTAGGTTCTCCGTCAGCTTCCCTCTCTCTGCGCATCTTTT 1380
Db 1321 AAAGATCAACAAGCTAGGTTCTCCGTCAGCTTCCCTCTCTCTGCGCATCTTTT 1380
Qy 1381 CGTCCACCAACATG 1394
Db 1381 CGTCCACCAACATG 1394

RESULT 2
US-10-713-381-1
; Sequence 1, Application us/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNANT, CARL W.
; APPLICANT: HUFPMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/680,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-1

Query Match 99.7%; Score 1389.2; DB 20; Length 1394;
Best Local Similarity 99.8%; Pred. No. 1.2e-293;
Matches 1391; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CCATGCTGCTCTATGAAAAAGATGAGTCAATGTCCTATATCCGTTTCTTAGAGTCC 60
Db 1 CCATGCTGCTCTATGAAAAAGATGAGTCAATGTCCTATATCCGTTTCTTAGAGTCC 60
Qy 61 CTTCTCTGCTTATTAAGTCACTGAATCGGGGTTTACAAAAAATTCCAGCGGTCATGAT 120
Db 61 CTTCTCTGCTTATTAAGTCACTGAATCGGGGTTTACAAAAAATTCCAGCGGTCATGAT 120
Qy 121 CTCGATGTTCCACTTCCTCCACCTCGGTCGATCATTTCTTGAGATCGGTCGATCCAT 180
Db 121 CTCGATGTTCCACTTCCTCCACCTCGGTCGATCATTTCTTGAGATCGGTCGATCCAT 180
Qy 181 CTGACCGAGGCCCATGACACCTTTGCGAGACCCCATCAAGGCGCTTTCGATGCGCA 240
Db 181 CTGACCGAGGCCCATGACACCTTTGCGAGACCCCATCAAGGCGCTTTCGATGCGCA 240
Qy 241 CGAGAGTATCGGGTCGTGTGATCCAGGGGATATATGTCCTCCCAATGCTCACTTAA 300
Db 241 CGAGAGTATCGGGTCGTGTGATCCAGGGGATATATGTCCTCCCAATGCTCACTTAA 300
Qy 301 TTATATTTCTTATGATATTTATTTTGAATAAATAAACAATTATACCTTTTGTGA 360
Db 301 TTATATTTCTTATGATATTTATTTTGAATAAATAAACAATTATACCTTTTGTGA 360
```

Db 301 TTATATCTTGGATATATTTATTTTGGAAAAATACAACTTATCTTTGGTGA 360
Qy 361 GGGCTTCAGATGATTTTGGCTTAGGGCCAGAAATGCGAGACGACCATGTCTAGT 420
Db 361 GGGCTTCAGATGATTTTGGCTTAGGGCCAGAAATGCGAGACGACCATGTCTAGT 420
Qy 421 TCGACTATGGGCTACCCGAGAACAGATTTAAAAATAACAAAGTAACTAATCCACT 480
Db 421 TCGACTATGGGCTACCCGAGAACAGATTTAAAAATAACAAAGTAACTAATCCACT 480
Qy 481 CGAAGCTATCATGATATGTTTAAAGAAATCTATTTAAACACGATCCTTTAAAAA 540
Db 481 CGAAGCTATCATGATATGTTTAAAGAAATCTATTTAAACACGATCCTTTAAAAA 540
Qy 541 CAAGCATATTTGCAAGAGACAAATTAATGTTACAGTTTCAACATCTAAGACGACAA 600
Db 541 CAAGCATATTTGCAAGAGACAAATTAATGTTACAGTTTCAACATCTAAGACGACAA 600
Qy 601 TTATATCGAAAGGTAAGCTATGACGTTACATTTTCTTTTCAATCTGTATTTGTT 660
Db 601 TTATATCGAAAGGTAAGCTATGACGTTACATTTTCTTTTCAATCTGTATTTGTT 660
Qy 661 ATTGTTTTATATACATTTTCTTCTTACATAGATGATTTTCTTCCGATTTTATAA 720
Db 661 ATTGTTTTATATACATTTTCTTCTTACATAGATGATTTTCTTCCGATTTTATAA 720
Qy 721 ATGACTATAAAGTCAATTTTATATAAGACGACATGCTAGATTTCTGTTCAAAAATC 780
Db 721 ATGACTATAAAGTCAATTTTATATAAGACGACATGCTAGATTTCTGTTCAAAAATC 780
Qy 781 TTCTCATTTTAAAGCTAGTTTGGCAACCTGTTTCTTTCAAGAAATTTGATTTT 840
Db 781 TTCTCATTTTAAAGCTAGTTTGGCAACCTGTTTCTTTCAAGAAATTTGATTTT 840
Qy 841 TTCAAAAAAATAGTTTATTTTCTTTTAAATAGAAACATTTAGAAAAATAGAT 900
Db 841 TTCAAAAAAATAGTTTATTTTCTTTTAAATAGAAACATTTAGAAAAATAGAT 900
Qy 901 TGGCAGACTAGCCCTAGAAATGTTTCCCAATTAATTAACAATCACTGTATTAATTTG 960
Db 901 TGGCAGACTAGCCCTAGAAATGTTTCCCAATTAATTAACAATCACTGTATTAATTTG 960
Qy 961 GCGAGCCCCATTAATTTTAAACCGAAATCGAATCGAGCGAAACCAATCTGACCTAT 1020
Db 961 GCGAGCCCCATTAATTTTAAACCGAAATCGAATCGAAGCAACCAATCTGACCTAT 1020
Qy 1021 TTCTAGATTAATTAAGAGAGAGAGAGAGAGAAATCAGTTTAAAGTCAATTTGCC 1080
Db 1021 TTCTAGATTAATTAAGAGAGAGAGAGAGAGAAATCAGTTTAAAGTCAATTTGCC 1080
Qy 1081 TGAGATGTGGGCTTTGGCAACGATAGCACCGTATCATAGTGTGCTTACGTCA 1140
Db 1081 TGAGATGTGGGCTTTGGCAACGATAGCACCGTATCATAGTGTGCTTACGTCA 1140
Qy 1141 GGTTCGGCAGCTCTGCTGATCTCAGATGGCATCTACTGCTTGTCAACCGTTCGTC 1200
Db 1141 GGTTCGGCAGCTCTGCTGATCTCAGATGGCATCTACTGCTTGTCAACCGTTCGTC 1200
Qy 1201 TTGTTCCATGCTCAAGACCTTGCCTATTTGGAACCAAGAGATACCTACCAACAT 1260
Db 1201 TTGTTCCATGCTCAAGACCTTGCCTATTTGGAACCAAGAGATACCTACCAACAT 1260
Qy 1261 CCATCTTACTAGCACTTCATGCAACGCAACATATGTTTCCGAAACCAATCAT 1320
Db 1261 CCATCTTACTAGCACTTCATGCAACGCAACATATGTTTCCGAAACCAATCAT 1320
Qy 1321 AAAGATCAACAAGCTAGCTTCTCCGCTAGCTTCTGCTCTGCTGCGATCTTTT 1380
Db 1321 AAAGATCAACAAGCTAGCTTCTCCGCTAGCTTCTGCTCTGCTGCGATCTTTT 1380
Qy 1381 CGTCACACACCATG 1394
Db 1381 CGTCACACACCATG 1394

Db 1381 CGTCACACACCATG 1394
RESULT 3
US-10-713-381-9
; Sequence 9, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 255
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-9
Query Match 17.5%; Score 243.6; DB 20; Length 255;
Best Local Similarity 98.4%; Pred. No. 2.9e-43;
Matches 246; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1145 CCGCAGCTCTGCTGATCTCAATGCGATGCTTGTCAACCGTTCGTTGT 1204
Db 5 CCGCGATCCCGTGTATCTCAATGCGATGCTTGTCAACCGTTCGTTGT 64
Qy 1205 TCCATGCTCAAGCTTGTCTTGTGAACGAAGGATACCTACCTCCAAACATCAT 1264
Db 65 TCCATGCTCAAGCTTGTCTTGTGAACGAAGGATACCTACCTCCAAACATCAT 124
Qy 1265 CTACTCATGCACTTCATGCAACGCAACATATGTTTCTGGAACAGATCTATTAAG 1324
Db 125 CTACTCATGCACTTCATGCAACGCAACATATGTTTCTGGAACAGATCTATTAAG 184
Qy 1325 ATCACAAGCTAGCGTTTCCCGCTAGCTTCTCTCTCTCTGCGATCTTTTCTGTC 1384
Db 185 ATCACAAGCTAGCGTTTCCCGCTAGCTTCTCTCTCTCTGCGATCTTTTCTGTC 244
Qy 1385 CACCAACATG 1394
Db 245 CACCAACATG 254
RESULT 4
US-10-713-381-3
; Sequence 3, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3

/ LENGTH: 158
/ TYPE: DNA
/ ORGANISM: Zea mays
US-10-713-381-3

Query Match 10.5%; Score 146; DB 20; Length 158;
Best Local Similarity 99.4%; Pred. No. 5.2e-22;
Matches 157; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1155 CGGTCATCCACATGGCATCTACTACGCTGTTCACACGCTGTC-TTGTTCATCGTC 1213
DB 1 CGGTATCTCACATGCGCATCTACTACGCTGTTCACACGCTGTCGTTCATCGTC 60

QY 1214 CAAGCCTGCTATCTTCTGGAACCAAGAGATACCTACTCCCAACAATCCATCTACTCAT 1273
DB 61 CAAGCCTTGCTATCTTCTGGAACCAAGAGATACCTACTCCCAACAATCCATCTACTCAT 120

QY 1274 GCACCTTCATGCAACACGCGACATATGTTTCTCTGAC 1311
DB 121 GCACCTTCATGCAACACGCGACATATGTTTCTCTGAC 158

RESULT 5
US-10-425-115-134230

/ Sequence 134230, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:

/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(53222)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 134230
/ LENGTH: 1261
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MRT4577_53903C.1
US-10-425-115-134230

Query Match 5.8%; Score 80.8; DB 20; Length 1261;
Best Local Similarity 78.2%; Pred. No. 2.7e-07;
Matches 97; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 792 TTAAAGCTAGTTTGGCAACCCGTTCTTCAAGAAATTTGATTTTCAAAAAA 851
DB 999 TCTTAAGGCTAGTTTGAAGACCTTATTTTAAAGATTTTCTTTTAAAGTAA 1058

QY 852 TTAGTTATTTTCTCTTTTAAATGAAAACCTTGAAGAAATAGAGTTGCCAGACTG 911
DB 1059 TTAGTTATTTTCTCTTTGAGAAATGAAATCTTTGAGAAATGAGTTGCTAAACTAG 1118

QY 912 CCT 915
DB 1119 CCT 1122

RESULT 6
US-10-739-930-1857

/ Sequence 1857, Application US/10739930
/ Publication No. US20040216190A1
/ GENERAL INFORMATION:

/ APPLICANT: Kovalic, David K.
/ TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
/ FILE REFERENCE: 38-21(53377)B
/ CURRENT APPLICATION NUMBER: US/10/739,930
/ CURRENT FILING DATE: 2003-12-18

/ NUMBER OF SEQ ID NOS: 11088
/ SEQ ID NO 1857
/ LENGTH: 1663
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: ZEAMA-23APR03-CLUSTER1429_8
US-10-739-930-1857

Query Match 5.5%; Score 76.2; DB 20; Length 1663;
Best Local Similarity 96.3%; Pred. No. 3.1e-06;
Matches 78; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1314 ATCTATTAAGATCAACAAGCTAGCTTCCGCTAGTCTCCCTCTCTGCGCA 1373
DB 2 ATCTATTAAGATCAACAAGCTAGCTTCCGCTAGTCTCCCTCTCTGCGCA 61

QY 1374 TCTTTTGTCTCACCACCATG 1394
DB 62 TCTTTTGTCTCACCACCATG 82

RESULT 7
US-10-425-114-23340

/ Sequence 23340, Application US/10425114
/ Publication No. US2004003488A1
/ GENERAL INFORMATION:

/ APPLICANT: Liu, Jingdong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E
/ APPLICANT: Tabaska, Jack E
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(53313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 23340
/ LENGTH: 320
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: LIB3595-039-G8_FLI
US-10-425-114-23340

Query Match 5.3%; Score 74; DB 18; Length 320;
Best Local Similarity 73.1%; Pred. No. 4.2e-06;
Matches 95; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 791 TTAAAGCTAGTTTGGCAACCCGTTCTTCAAGAAATTTGATTTTCAAAAAA 850
DB 62 TTAAAGGCTAGTTTGAAGACCTTATTTTAAAGATTTTCTTTTAAAGTAA 121

QY 851 ATTAGTTATTTTCTCTTTTAAATGAAAACCTTGAAGAAATAGAGTTGCCAGACTA 910
DB 122 ATTAGTTATTTTCTCTTTGAGAAATGAAATCTTGGAAATAGAGTTCTAAACGA 181

QY 911 GCCCTGAAT 920
DB 182 GCCCTGAAT 191

RESULT 8
US-10-425-114-16264

/ Sequence 16264, Application US/10425114
/ Publication No. US2004003488A1
/ GENERAL INFORMATION:

/ APPLICANT: Liu, Jingdong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E

	Query Match	Similarity	5.2%	Score 73	DB 20	Length 126
	Best Local	Similarity	69.1%	Pred. No. 1.4e-05		
	Matches	114	Conservative	0	Mismatches 50	Indels 1; Gaps 1
Qy	756	TGTCGTAGATTCCTCGTCAAAATCTTTCTGATTTTCTTAAAGAGCTAGTTGGCAACCT	815			
Db	1090	TGAAGTTGGTGGTGTGCTGGATTTTCACTTAAGTAAAGACTAATTTGGGAACCA	1149			
Qy	816	GTTTCTTTCAAGAAATTTTGATTTTTCACAAAAAATAGATTATTTCTCTTTATAAA	875			
Db	1150	ATTT-TTCCACCTGGATTTTTCATTTCTTAAGAAAAATAGTTCATTTTCCCTTGAGAAA	1208			
Qy	876	TAGAAAACTTAGAAAAATAGTTGGCAAGCTAGCCCTGAAT	920			
Db	1209	TAGGAATCCCTTAGAAAAAATAGTTTCCAAACTAGCCCTTAAT	1253			

Query Match	5.2%;	Score 72.8;	DB 20;	Length 2729;
Best Local Similarity	60.7%;	Pred. No. 2.2e-05;		
Matches 136;	Conservative 0;	Mismatches 87;	Indels 1;	Gaps 1;

```
Qy 792 TTTAAGAGCTAGTTGGCAACCTGTTCTTTCAAGAAATTTGATTTTTCACAAAAA 851
    |||||
Db 2523 TTTAAGGCTAGTTAGAA-CCTCATTTTCTAGAGATTTCTATTTTCCAGAAAA 2465
Qy 852 TTAAGTTATTTCTCTTTATAAATAGAAAACTTAGAAAAATAGAGTTGCCAGACTAG 911
    |||||
Db 2464 ATATTTTCATTTCCCTTGAAAAATAGAAATACCTTGAAAAATAGAGTTCCAACTAG 2405
Qy 912 CCTAGAAATGTTTCCCAATTAATTAACAATGTAATTAATTTGGCCAGCCCAT 971
    |||||
Db 2404 CTTCAAAATTTCTCTCAAGATATTTCTTATAGAGATTAAGTACACATGGGAGCT 2345
Qy 972 AAATATTAAACCGAACTGAATCGAGGAAACCAATCTGA 1015
    |||||
Db 2344 TGAGTATCATGAGCTAAATGTAAAGGGTAAATCAATGTTA 2301
```

```
RESULT 12
US-10-425-115-51470/c
; Sequence 51470, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 51470
; LENGTH: 1203
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1203)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_146936C.1
US-10-425-115-51470
```

```
Query Match 5.2%; Score 72.4; DB 20; Length 1203;
Best Local Similarity 76.1%; Pred. No. 1.8e-05;
Matches 102; Conservative 0; Mismatches 31; Indels 1; Gaps 1;
```

```
Qy 787 ATTTTAAAGAGCTAGTTGGCAACCTGTTCTTTCAAGAAATTTGATTTTTCAAA 846
    |||||
Db 1189 ACTTTAAGGCTAGTGGGGA-CATATTTTCCAAAAAGATTTCTATTTTCTTAA 1131
Qy 847 AAAAATTAAGTTATTTCTCTTTATAAATAGAAAACTTAGAAAAATAGAGTTGCCAG 906
    |||||
Db 1130 GAAAAATTAATTTTCTCTTGAAAAATAAATCCGTAGAAAAATAGAGTTCCAA 1071
Qy 907 ACTAGCCCTAGAA 920
    |||||
Db 1070 ACTAGCTTAGAT 1057
```

```
RESULT 13
US-10-425-114-17816/c
; Sequence 17816, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
```

```
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 17816
; LENGTH: 928
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3075-032-A3_FLI
US-10-425-114-17816
```

```
Query Match 5.2%; Score 71.8; DB 18; Length 928;
Best Local Similarity 78.0%; Pred. No. 2.1e-05;
Matches 99; Conservative 0; Mismatches 27; Indels 1; Gaps 1;
```

```
Qy 790 TTTTAAGAGCTAGTTGGCAACCTGTTCTTTCAAGAAATTTGATTTTTCAAAAA 849
    |||||
Db 804 TTTTAAGAGCTGTTGGCA-CCTCATTTTCTAAGGATTTCTATTTTCTAAGAA 746
Qy 850 AATTAGTTATTTCTCTTTATAAATAGAAAACTTAGAAAAATAGAGTTGCCAGACT 909
    |||||
Db 745 AATTAGTTAATTTCTCTTAGAAAAATAAATGTCTTGAAAAATAGAGTTACCAAT 686
Qy 910 AGCCCTA 916
    |||||
Db 685 ACCCCTA 679
```

```
RESULT 14
US-10-425-115-142853/c
; Sequence 142853, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 142853
; LENGTH: 1123
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_61761C.1
US-10-425-115-142853
```

```
Query Match 5.2%; Score 71.8; DB 20; Length 1123;
Best Local Similarity 78.0%; Pred. No. 2.3e-05;
Matches 99; Conservative 0; Mismatches 27; Indels 1; Gaps 1;
```

```
Qy 790 TTTTAAGAGCTAGTTGGCAACCTGTTCTTTCAAGAAATTTGATTTTTCAAAAA 849
    |||||
Db 821 TTTTAAGAGCTGTTGGCA-CCTCATTTTCTAAGGATTTCTATTTTACTAAGAA 763
Qy 850 AATTAGTTATTTCTCTTTATAAATAGAAAACTTAGAAAAATAGAGTTGCCAGACT 909
    |||||
Db 762 AATTAGTTATTTCTCTTAGAAAAATAAATGTCTTGAAAAATAGAGTTACCAAT 703
Qy 910 AGCCCTA 916
    |||||
Db 702 ACCCCTA 696
```

```
RESULT 15
US-10-425-114-31957
; Sequence 31957, Application US/10425114
```

```
; Publication No. US20040034889A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 31957
; LENGTH: 2537
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73238H12_FLI
; US-10-425-114-31957

Query Match      5.2%; Score 71.8; DB 18; Length 2537;
Best Local Similarity 73.4%; Pred. No. 3.5e-05;
Matches 105; Conservative 0; Mismatches 37; Indels 1; Gaps 1;

QY      792 TTTAGAGCTAGTTGGCAACCTGTTCTTTCAAGAAATTTGATTTTTCAAAAAAA 851
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      1799 TCTTAGGGCTAGTTGGAACCT-TTTTCCCAAAAGATTTTCATTTTCMAAGAAAA 1857

QY      852 TTAGTTATTTCTCTTATAAATAGAAACCTAGAAAAATAGATTGCCAGACTAG 911
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      1858 TTAGTTCATTTTCTTGAAGAAATAGGAATCCCTTAAAAAATAGTGTTCMAACTAG 1917

QY      912 CCTAGAAATGTTTCCCAATAAA 934
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      1918 CCCTTAATGTTTTCATGAA 1940
```

Search completed: September 15, 2005, 20:45:53
Job time : 908.859 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 16:07:32 ; Search time 4347.69 Seconds
(without alignments)
12204.542 Million cell updates/sec

Title: US-10-713-381-2

Perfect score: 1394

Sequence: 1 cccttgctgcctcctcgaagaa.....cttttcgctccaccaccatg 1394

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_eat1:*
2: gb_eat2:*
3: gb_hlc:*
4: gb_eat3:*
5: gb_eat4:*
6: gb_eat5:*
7: gb_eat6:*
8: gb_g881:*
9: gb_g882:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	948.8	68.1	963	9	CC656933 OGMQ20TM
2	755.6	54.2	915	9	CG224225 OG1AG08TV
3	496.4	35.6	687	9	CC656939 OGMQ20TV
4	96	6.9	715	9	CG252571 CG4B05TC
5	92	6.6	967	9	CL235046 ZMMBB057
6	88	6.3	814	9	CG048704 PU1019TB
7	87.8	6.3	754	9	CG414922 ZMMBB029
8	86.8	6.2	950	8	CC439901 PUHRI15TB
9	86.6	6.2	652	8	CC384247 PUHOC7TB
10	86.6	6.2	797	8	CC400575 PUHUI61TD
11	86.6	6.2	820	8	CC400574 PUHUI61TB
12	85.4	6.1	471	9	CG103452 PUHRI9TB
13	85.2	6.1	765	9	CG082135 PUHRI12TD
14	85.2	6.1	781	9	CC630219 OGUCG53TV
15	85.2	6.1	815	9	CG349565 OG0F183TV
16	85.2	6.1	834	9	CC630210 OGUCG53TV
17	85	6.1	789	8	CC433618 PUHRI17TD
18	84.6	6.1	1092	8	CNS020K7
19	84	6.0	793	8	BZ816381
20	83.8	6.0	1078	9	CL997678 ZMMBH001
21	83.6	6.0	530	9	CG201774 PUICH24TB
22	83.6	6.0	722	9	CG333914 OG0AD14TV
23	83.6	6.0	722	9	CG333929 OG0AD14TV
24	83.6	6.0	781	9	CG034985 PUHRI68TB

25	83.6	6.0	861	8	BZ797976
26	83.6	6.0	865	8	CC430754
27	81.8	5.9	861	9	CG102052
28	81.8	5.9	947	8	CC435780
29	81.8	5.9	981	8	BZ784278
30	81.8	5.9	1017	9	CC620594
31	81.6	5.9	649	9	CC613918
32	81.6	5.9	733	8	BZ778636
33	81.6	5.9	925	8	CG071791
34	81.6	5.9	960	8	BZ676889
35	81.6	5.9	994	8	CC003943
36	81.6	5.9	1016	9	CL996481
37	81.6	5.9	1022	9	CL984151
38	81.4	5.8	999	8	CC385762
39	81.2	5.8	756	9	CG221653
40	81	5.8	1101	9	CNS00172
41	80.6	5.8	699	8	BZ96930
42	80.6	5.8	832	9	CC676387
43	80.6	5.8	839	9	CG246849
44	80.6	5.8	843	9	CG254365
45	80.6	5.8	852	8	CC385881

ALIGNMENTS

RESULT 1
LOCUS OGMQ20TM 963 bp DNA linear GSS 19-JUN-2003
DEFINITION genomic survey sequence.
ACCESSION CC656933
VERSION CC656933.1 GI:32060225
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 963)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T., Resnick,A., Fraser,C.M., Buddiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSs: OGMQ20TV
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@cigr.org
Seq primer: TR
Class: sheared ends.

FEATURES

source
1..963
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMMA0554D15"
/clone_1lb="ZM_0.7-1.5_KB"
/note="Vector: pBCK-; Site 1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"

ORIGIN

Query Match 68.1%; Score 948.8; DB 9; Length 963;
Best Local Similarity 99.7%; Pred. No. 1.5e-190;
Matches 961; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY 404 ACCAGCATGTCTAGTGTCTATTTGGCACTACCAAGAAAGATTAAATTAACC 463
DB 1 ACCAGCATGTCTAGTGTCTATTTGGCACTACCAAGAAAGATTAAATTAACC 60

```

QY      464 AAGTAATCAATCCACTCGAAGACTATCATGTAATGTTTAAAGAAACATCTATTTAAACC 523
      |||||
Db      61 AAGTAATCAATCCACTCGAAGACTATCATGTAATGTTTAAAGAAACATCTATTTAAACC 120
QY      524 AGCATCTCTTAAACAAAGCATATTTGCAAAAGACAAATTAATGTTTACAGTTTACAA 583
      |||||
Db      121 ACATCTCTTAAACAAAGCATATTTGCAAAAGACAAATTAATGTTTACAGTTTACAA 180
QY      584 CATCTAAGACGCAAAATTAATGCAAAAGTAGCTATGACGTTGAGATTTTCTTTTTC 643
      |||||
Db      181 CATCTAAGACGCAAAATTAATGCAAAAGTAGCTATGACGTTGAGATTTTCTTTTTC 240
QY      644 ATCTCTGTAATTTGTAATGTTTATATATACATTTTCTCTCTTCAATAGAGTATTT 703
      |||||
Db      241 ATCTCTGTAATTTGTAATGTTTATATATACATTTTCTCTCTTCAATAGAGTATTT 300
QY      704 TCTTCGATTTTAAATGACTATTAAGTCAATTTTATTAATTAAGACGCAATGTCGTAG 763
      |||||
Db      301 TCTTCGATTTTAAATGACTATTAAGTCAATTTTATTAATTAAGACGCAATGTCGTAG 360
QY      764 ATCTCTGTAATTTCTTCTGATTTTAAAGTCAATTTTAAAGTCAATTTTAAAGTCAATTT 823
      |||||
Db      361 ATCTCTGTAATTTCTTCTGATTTTAAAGTCAATTTTAAAGTCAATTTTAAAGTCAATTT 420
QY      824 CAAGAATTTTGAATTTTTCAAAAAATTAATGTTTATTTCTCTTATTAATTAAGAAAC 883
      |||||
Db      421 CAAGAATTTTGAATTTTTCAAAAAATTAATGTTTATTTCTCTTATTAATTAAGAAAC 480
QY      884 ACTTGAATAAATAGAGTGGCAGACTAGCCCTGAAATGTTTCCCAATAAATTAATCAATCA 943
      |||||
Db      481 ACTTGAATAAATAGAGTGGCAGACTAGCCCTGAAATGTTTCCCAATAAATTAATCAATCA 540
QY      944 CTGTGTATATTTTGGCCAGCCCAATAATTAATTAAGTCAATTTTAAAGTCAATTTTAAAGTCAATTT 1003
      |||||
Db      541 CTGTGTATATTTTGGCCAGCCCAATAATTAATTAAGTCAATTTTAAAGTCAATTTTAAAGTCAATTT 600
QY      1004 AACCAATCTGACCTATTTCTCTAGATTAATTAAGTCAATTTTAAAGTCAATTTTAAAGTCAATTT 1063
      |||||
Db      601 AACCAATCTGACCTATTTCTCTAGATTAATTAAGTCAATTTTAAAGTCAATTTTAAAGTCAATTT 660
QY      1064 TTTTAAAGTCAATTTCTCTGAGATGTCGGTGGCAACGATAGCCGATATCATAGCT 1123
      |||||
Db      661 TTTTAAAGTCAATTTCTCTGAGATGTCGGTGGCAACGATAGCCGATATCATAGCT 720
QY      1124 CATAGGAGCTAGTCAAGTGGGCAAGCTCTGATCTCATGAGGCAATCATATCATAGC 1183
      |||||
Db      721 CATAGGAGCTAGTCAAGTGGGCAAGCTCTGATCTCATGAGGCAATCATATCATAGC 780
QY      1184 TTGTTCAACCGTCTGCTTGTTCATCTGTCGAAGCCCTGCTATTTCTGAACCAAGAGAT 1243
      |||||
Db      781 TTGTTCAACCGTCTGCTTGTTCATCTGTCGAAGCCCTGCTATTTCTGAACCAAGAGAT 840
QY      1244 ACTTACTCCCAACCAATCTTACTCATGCACTTCCATGCAACAGCAGATATGTT 1303
      |||||
Db      841 ACTTACTCCCAACCAATCTTACTCATGCACTTCCATGCAACAGCAGATATGTT 900
QY      1304 TCTGTAACGATATTAAGATCAACAAGCTAGAGGTTCTCCGCTAGGTTCCCTCTCT 1363
      |||||
Db      901 TCTGTAACGATATTAAGATCAACAAGCTAGAGGTTCTCCGCTAGGTTCCCTCTCT 959
QY      1364 CCTC 1367
      |||||
Db      960 CCTC 963

```

```

RESULT 2
LOCUS   CG224225
DEFINITION OGIAG08TV_ZM_0_7_1_5_KB_Zea_mays_genomic_clone_ZMMBma0716b15,
ACCESSION CG224225
VERSION   CG224225.1 GI:34124113

```

```

KEYWORDS GSS.
SOURCE   Zea mays
ORGANISM Zea mays
REFERENCE 1 (bases 1 to 915)
AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Uteerback,T.,
          Resnick,A., Fraser,C.M., Buddiman,M.A., Hedell,J.A., Rohlfing,T.,
          Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
          Consortium for Maize Genomics
          Unpublished (2002)
          Other GSSes: OGIAG08TV
          Contact: Cathy Whitelaw
          TIGR
          9712 Medical Center Drive, Rockville, MD 20850, USA
          Tel: 301-838-5843
          Fax: 301-838-0208
          Email: whitelaw@cigr.org
          Seq primer: TP
          Class: sheared ends.
FEATURES
    source          location/Qualifiers
                    1..915
                    /organism="Zea mays"
                    /mol_type="genomic DNA"
                    /strain="B73"
                    /db_xref="taxon:4577"
                    /clone="ZMMBma0716b15"
                    /clone_lib="ZM_0_7_1_5_KB"
                    /note="vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
                    methylation filtered genomic DNA library"
ORIGIN
Query Match          54.2%; Score 755.6; DB 9; Length 915;
Best Local Similarity 99.4%; Pred. No. 1,3e-149;
Matches 769; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
QY      621 TGACGTTGAGATTTTCTTTTTCATTTCTGTTATTTGTTATTTTAAATACATTTT 680
      |||||
Db      1 TGACGTTGAGATTTTCTTTTTC-TTCTGTTATTTGTTATTTGTTTATATACATTTT 59
QY      681 CTTCCTTCAATTAAGATGATTTTCTCCGATTTTATTAATTAAGTCAATTTT 740
      |||||
Db      60 CTTCCTTCAATTAAGATGATTTTCTCCGATTTTATTAATTAAGTCAATTTT 119
QY      741 AATTAAGACGACGATGTCGTATTTCTGTCAAAATCTTTGATTTTAAAGGC 800
      |||||
Db      120 AATTAAGACGACGATGTCGTATTTCTGTCAAAATCTTTGATTTTAAAGGC 179
QY      801 TAGTTGGCAACCGTCTTCTTCAAGAAATTTGATTTTCAAAAAAATTAAGTTAT 860
      |||||
Db      180 TAGTTGGCAACCGTCTTCTTCAAGAAATTTGATTTTCAAAAAAATTAAGTTAT 239
QY      861 TTTCTCTTATTAATTAAGAAACACTTAAGAAATTAAGTTGCCAGTACGCCCTAGAT 920
      |||||
Db      240 TTTCTCTTATTAATTAAGAAACACTTAAGAAATTAAGTTGCCAGTACGCCCTAGAT 299
QY      921 GTTTTCCCAATTAATTAAGTCACTGTAATTAATTTGGCAGCCCATTAATTT 980
      |||||
Db      300 GTTTTCCCAATTAATTAAGTCACTGTAATTAATTTGGCAGCCCATTAATTT 359
QY      981 AAACGAAACTGAATTCAGGCAACCAAACTGAGTATTTCTGATTAATTAAGTAAAG 1040
      |||||
Db      360 AAACGAAACTGAATTCAGGCAACCAAACTGAGTATTTCTGATTAATTAAGTAAAG 419
QY      1041 GGAGAGAGAGAGAAATCAATGTTTAAGTCAATTTGCCAGATGTCGGTTGGCAA 1100
      |||||
Db      420 GGAGAGAGAGAGAAATCAATGTTTAAGTCAATTTGCCAGATGTCGGTTGGCAA 479
QY      1101 CGATAGCAACCGTATCATAGTCAATGTCCTGATGCTGAGTGGGAGCTCTGTC 1160
      |||||
Db      480 CGATAGCAACCGTATCATAGTCAATGTCCTGATGCTGAGTGGGAGCTCTGTC 539

```

QY 1161 ATCTACATGSCATACACTAGCTGTTGTTCAACGGTTCGTGTTGTTCCATGCTCAAGCCT 1220
|||||
Db 540 ATCTCATATGSCATACACTAGCTGTTGTTCAACGGTTCGTGTTGTTCCATGCTCAAGCCT 599
|||||
QY 1221 TGGCTATTTCTGAACCAAGAGATACCTACTCCCAACATCATCTTACTCATGCAACT 1280
|||||
Db 600 TGGCTATTTCTGAACCAAGAGATACCTACTCCCAACATCATCTTACTCATGCAACT 659
|||||
QY 1281 CCATGCAAAACGACATATGTTTCTGGAACAGATCTATTAAAGATACAAACGCTAGCG 1340
|||||
Db 660 CCATGCAAAACGACATATGTTTCTGGAACAGATCTATTAAAGATACAAACGCTAGCG 719
|||||
QY 1341 TTCTCCGCTAGGCTCCCTCTCTCTGCGGATCTTTTTCGTCCACCAACGATG 1394
|||||
Db 720 TTCTCCGCTAGGCTCCCTCTCTCTGCGGATCTTTTTCGTCCACCAACGATG 773
|||||
RESULT 3
CC656939 687 bp DNA linear GSS 19-JUN-2003
LOCUS CGMD020TV_ZM_0.7_1.5_KB_Zea_mays_genomic_clone_ZM54D15,
DEFINITION genomic survey sequence.
ACCESSION CC656939
VERSION CC656939.1 GI:32060231
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 687)
White, C.A., Quackenbush, J., Van Aken, S., Uterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
Other_GSSs: CGMD020TM
COMMENT Contact: Cathy White, Tel:
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitec@tigr.org
Seq primer: TF
Class: sheared ends.
FEATURES
source 1..687
Location/Qualifiers
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM54D15"
/clone_1lb="ZM_0.7_1.5_KB"
/note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"
ORIGIN
Query Match 35.6%; Score 496.4; DB 9; Length 687;
Best Local Similarity 98.8%; Pred. No. 1.1e-94;
Matches 500; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 889 GAAAAATAGAGTTGCCAGACTAGCCCTAGATGTTTCCCAATTAATTACATCTACTG 948
|||||
Db 687 GAAAAATAGAGTTGCCAGACTAGCCCTAGATGTTTCCCAATTAATTACATCTACTG 628
|||||
QY 949 TATAATATTATTTGGCGAGCCCATTAATTAATTAAACGAATCGAATCGAAGAACCA 1008
|||||
Db 627 TATAATATTATTTGGCGAGCCCATTAATTAATTAAACGAATCGAATCGAAGAACCA 568
|||||
QY 1009 AATCTGAGTAATTTCTTGAATTAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1068
|||||
Db 567 AATCTGAGTAATTTCTTGAATTAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 508
|||||

QY 1069 AGTCATTTGTCCTGAGATGAGGCTTTGGCAACGATAGGACCGATATCATAGCTATAG 1128
|||||
Db 507 AGTCATTTGTCCTGAGATGAGGCTTTGGCAACGATAGGACCGATATCATAGCTATAG 448
|||||
QY 1129 GTGCTTAGCTAGAGTTGCGAGCTCTGTGATCTACATCTGATGATGATGATGATGAT 1188
|||||
Db 447 GTGCTTAGCTAGAGTTGCGAGCTCTGTGATCTACATCTGATGATGATGATGATGAT 388
|||||
QY 1189 CAACGCTTGTCTTTGTTCCATGCTCAAGCTTGTCTATTTGTAACCAAGAGATACCTA 1248
|||||
Db 387 CAACGCTTGTCTTTGTTCCATGCTCAAGCTTGTCTATTTGTAACCAAGAGATACCTA 328
|||||
QY 1249 CTCGCAAAACATCATCTTACTCATGCACTTCATGCAAAACGACGATATGTTCTG 1308
|||||
Db 327 CTCGCAAAACATCATCTTACTCATGCACTTCATGCAAAACGACGATATGTTCTG 268
|||||
QY 1309 AACAGATCTATTAAAGATCAACAACGCTTCTCCGCTAGCTTCCCTCTCTCTCT 1368
|||||
Db 267 AACAGATCTATTAAAGATCAACAACGCTTCTCCGCTAGCTTCCCTCTCTCTCTCT 208
|||||
QY 1369 GCCGATCTTTTTCGTCCACCAACGATG 1394
|||||
Db 207 GCCGATCTTTTTCGTCCACCAACGATG 182
|||||
RESULT 4
CG252571 715 bp DNA linear GSS 25-AUG-2003
LOCUS CG4B057C_ZM_0.7_1.5_KB_Zea_mays_genomic_clone_ZM54D10,
DEFINITION genomic survey sequence.
ACCESSION CG252571
VERSION CG252571.1 GI:34154661
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 715)
White, C.A., Quackenbush, J., Van Aken, S., Uterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
COMMENT Contact: Cathy White, Tel:
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitec@tigr.org
Seq primer: TF
Class: sheared ends.
FEATURES
source 1..715
Location/Qualifiers
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM54D10"
/clone_1lb="ZM_0.7_1.5_KB"
/note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"
ORIGIN
Query Match 6.9%; Score 96; DB 9; Length 715;
Best Local Similarity 80.1%; Pred. No. 8.3e-10;
Matches 125; Conservative 0; Mismatches 30; Indels 1; Gaps 1;
QY 135 TCTCCACCTCGCTTGACATTTCTTGATGTCGTTGCTCCATCTGACGAGGCCCA 194
|||||
Db 356 TCACACACCGCTCTTGTCCTTTCTTGATGTCGCGGTTCCGCTGATGAGGCCCA 297
|||||
QY 195 TCAGACACCTTTTCGGAGACCAATCAAGGCTTTTCGATGACCAAGACGATAGCG 254
|||||

Db 296 TAAGGACACGTGTCGGAGAACCCATTAAAGCCCTTTCGATGCGCCAGGACAT-CGGGG 238
QY 255 TCGTGTGATCCAGAGGATATATGTCCCCCAACATC 290
Db 237 TCGTGTGACCCGGGGATATATATCCCAACACC 202

RESULT 5
CL235046 967 bp DNA linear GSS 15-JAN-2004
LOCUS ZMBBB0575001r ZMBBB (HindIII) Zea mays genomic clone
DEFINITION ZMBBB0575001 3', genomic survey sequence.
ACCESSION CL235046
VERSION CL235046.1 GI:40891729
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 967)
AUTHORS Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C.,
Zohovetz,V., Fuks,G., Yu,Y., Wing,R. and Messing,J.
JOURNAL Sequencing of the maize genome at PGIR (2003c)
COMMENT Unpublished (2003)
Contact: Bharti,A.K.
Dr.Joachim Messing's lab
The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
University
190 Frelinghuysen Road, Piscataway, NJ 08854, USA
Tel: 732 445 3801
Fax: 732 445 5735
Email: bharti@waksman.rutgers.edu
Seq primer: SP6
Class: BAC ends
High quality sequence start: 64.
Location/Qualifiers
1. 967
/organism="Zea mays"
/mol_type="genomic DNA"
/culivar="B73"
/db_xref="taxon:4577"
/clone="ZMBBB0575001"
/lab_host="E. coli DH10B"
/clone_lib="ZMBBB (HindIII)"
/note="Vector: PCUG1; Site_1: HindIII; Site_2: HindIII"

ORIGIN

Query Match 6.6%; Score 92; DB 9; Length 967;
Best Local Similarity 78.6%; Pred. No. 5.9e-09;
Matches 110; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 150 TGCACATTTCTTGAGATGCGTGTGTCATCTGACCGAGCCCATCAGACACCTTTCCG 209
Db 827 TGGCGCTTCTTGAGATCCGCGCGTTCCTCGTGAATGAGCCCATCCAGACACTTTTGG 886
QY 210 GACACCCATCAAGGCGCTTTCGATGCGCCACGAGACGTATCGGTCGTGTATCCAGG 269
Db 887 GTGGCCCATCAAGTGCCTTTCGATGTCACGCGAGATGATCCGGGCCGCTGTGACCTGGG 946
QY 270 GGATATATGTCGCCCAACAT 289
Db 947 GGATATCTGTCCCCCATAGT 966

RESULT 6
CG048704 814 bp DNA linear GSS 19-AUG-2003
LOCUS PU1L019TB ZM_0_6_1_0_KB Zea mays genomic clone ZMBBB0611C13,
DEFINITION genomic survey sequence.
ACCESSION CG048704
VERSION CG048704.1 GI:33920884

KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 814)
AUTHORS Whitelew,C.A., Quackenbush,J., Van Aken,S., Uteerback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
JOURNAL Maize Genomics Consortium
COMMENT Unpublished (2003)
Other GSSes: PU1L019TB
Contact: Cathy Whitelew
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelew@cigr.org
Seq primer: TR
Class: sheared ends.
Location/Qualifiers
1. 814
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMBBB0611C13"
/clone_lib="ZM_0_6_1_0_KB"
/note="Vector: pCR4-toPO; Site_1: EcoRI; 0.6-1.0 kb high
cot selected genomic DNA library"

ORIGIN

Query Match 6.3%; Score 89; DB 9; Length 814;
Best Local Similarity 62.7%; Pred. No. 4.1e-08;
Matches 153; Conservative 0; Mismatches 90; Indels 1; Gaps 1;

QY 713 TTTATTAATGACTTAAAGTCATTTTATATAGACGACGATGTCGTATTCGTT 772
Db 487 TTTAGAGACGTTGTGATGATGAGAAAGATATAGAAATATATCTTTAGGAATGTTA 546
QY 773 CAAAAATCTTCTGATTTTATTTTAAAGCTAGTTTGCAACCTGTCTTTCAAAGATT 832
Db 547 TAAAGGACAAAGAAATTTCTTAAGAGCTAGTTGTATCCCATTTTTCAAAGATT 606
QY 833 TTGATTTTTCAAA-AAAATTAGTTTATTTCTTTATATAAAGAAACATTAGAA 891
Db 607 TTCATTTCACAAAGGAAATATGTTTATTTCCCTTTGTAAATGTAATCCTCGAA 666
QY 892 AAATAGATGTCAGACTAGCCCTAGAAATGTTTCCCAATAATTAACAATCACTGTGAT 951
Db 667 AAATAGATTTCAAACTAGCCCTTAAGGTTATATCCCATTAATCCCTATTCATTTCT 726
QY 952 AATT 955
Db 727 ATTT 730

RESULT 7
CG414922 754 bp DNA linear GSS 08-SEP-2003
LOCUS ZMBBB0290L09.r ZMBBB Zea mays genomic clone ZMBBB0290L09 3',
DEFINITION genomic survey sequence.
ACCESSION CG414922
VERSION CG414922.1 GI:34505144
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 754)
AUTHORS Yu,Y., Kim,H.R., Hatfield,J., Soderlund,C., Bharti,A.K., Messing,J.

TITLE and Wing, R.
JOURNAL Sequencing of the maize genome
COMMENT Unpublished (2003)
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088 USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: <http://genome.arizona.edu>
PCR Primers
FORWARD: T7
REVERSE: M13r
BACKWARD: M13r
Plate: 0290 row: L column: 09
Seq primer: M13r
Class: BAC ends.

FEATURES

source Location/Qualifiers
1..754
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ZM58B0290L09"
/lab_host="DH108"
/clone_1lb="ZM58Bb"
/note="Vector: pBeloBAC11, Site_1: HindIII, Site_2: HindIII; Zea mays L. ssp. mays"

ORIGIN

Query Match 6.3%; Score 87.8; DB 9; Length 754;
Best Local Similarity 71.2%; Pred. No. 4.5e-08;
Matches 116; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 779 TCTTTCGATTTTAAAGAGCTAGTTGGCAACCCCTGTTCTTCAAGATTTGATT 838
DB 680 TCTAAATTTATTCCTAAAGCTAGTTGAAAGCAATTTTCCAGAGATTTTATTT 621
QY 839 TTTTCAAAAAAATTAGTTATTTTCTTTATTAATAAGAAACACTTGAAAAATAGA 898
DB 620 TTTCTAAATTAATAATTAATTTCTTGAATAAATAGAAATCAGATGAAAAATATA 561
QY 899 GTTCCGAGCTAGCCCTAGATGTTTCCCAATAATTACAT 941
DB 560 GTTCTTAACCTAAGCTTAATAAATATTTCTTGCTCATTTACAGT 518

RESULT 8
CC439901/c 950 bp DNA linear GSS 20-MAY-2003
LOCUS PUHRV15TB.ZM.0.6.1.0_KB Zea mays genomic clone ZMBR7A519D06,
DEFINITION genomic survey sequence.
ACCESSION CC439901
VERSION CC439901.1 GI:30940477
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 950)
White, C.A., Quackenbush, J., Van Aken, S., Uterback, T.,
Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and
Benneken, J.
Maize Genomics Consortium
Unpublished (2003)
Other_GSSs: PUHRV15TD
Contact: Cathy White, whitelaw@ligr.org
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@ligr.org

Seq primer: TR
Class: sheared ends.
Location/Qualifiers
1..950
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM58B0290L09"
/note="Vector: pCR4-TOPO, Site_1: EcoRI; 0.6-1.0 kb high
COT selected genomic DNA library"

ORIGIN

Query Match 6.2%; Score 86.8; DB 8; Length 950;
Best Local Similarity 72.4%; Pred. No. 7.5e-08;
Matches 126; Conservative 0; Mismatches 47; Indels 1; Gaps 1;

QY 787 ATTTTAAAGAGCTGTTGGCAACCCCTGTTCTTCAAGATTTGATTTCATA 846
DB 241 ATTTTCTAATGAGCTAGTTGGCAACCTTATTT-TCACGGGATTTTATTTTCAAT 183
QY 847 AAAAATTAGTTATTTTCTTTATTAATAAGAAACACTTAGAAAAATAGAGTCCAG 906
DB 182 GAAATTTAGTTATTTTCTCTTGAGAAATATTAATCAGTTGAGAAATATAGTTCCAA 123
QY 907 ACTAGCCCTAGATGTTTCCCAATAATTACATCAGTGTATATTTTG 960
DB 122 ACTAGCCCTAAGATTTATTAATAATAGAAAAAATGCTATATTTTCATTG 69

RESULT 9

CC384247 652 bp DNA linear GSS 19-MAY-2003
LOCUS PUHOC67TB.ZM.0.6.1.0_KB Zea mays genomic clone ZMBR7A495L13,
DEFINITION genomic survey sequence.
ACCESSION CC384247
VERSION CC384247.1 GI:30864143
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 652)
White, C.A., Quackenbush, J., Van Aken, S., Uterback, T.,
Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and
Benneken, J.
Maize Genomics Consortium
Unpublished (2003)
Other_GSSs: PUHOC67TD
Contact: Cathy White, whitelaw@ligr.org
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@ligr.org
Seq primer: TR
Class: sheared ends.
Location/Qualifiers
1..652
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM58B0495L13"
/note="Vector: pCR4-TOPO, Site_1: EcoRI; 0.6-1.0 kb high
COT selected genomic DNA library"

ORIGIN
Query Match 6.2%; Score 86.6; DB 8; Length 652;
Best Local Similarity 63.9%; Pred. No. 8.1e-08;

Matches	131	Conservative	0	Mismatches	74	Indels	0	Gaps	0
QY	714	TTATAAAATGACTATPAAGTCATTTTATATATAAGACACGCATGTCGTAGATTCTGCTT							773
DB	341	TAAATTAATTTGTAGTACTGACCTACATCTTAAATCGACATATTAATTTAGTTAAATTAATGCTA							400
QY	774	AAAAATCTTTCGATTTTTTTTAAGAGCTAGTTTGCAACCCGTTCTTTTCAAAAGATT							833
DB	401	AGAACTTATTTATTTAATTAAGTAAAGCTAGTTTGCAACCTAATTTTTCTTAAGACATTTT							460
QY	834	TGATTTTTTCAAAAAAAATTAAGTTATTTTCTTTATATAAATAGAAAAACACTTGAAGA							893
DB	461	TATTTTTTCAAGTGAATTAATGTTATTTTTCATGCGAAAAATTAATTAATTTGGAAGA							520
QY	894	ATAGAGTTGCCAGACTAGCCCTAGA	918						
DB	521	ATTAGTTTCCAAACTAGCCCTAAA	545						
RESULT 10									
CC400575/c									
LOCUS									
DEFINITION	CC400575	797 bp	DNA	linear	GSS	19-MAY-2003			
ACCESSION	PURLHUTD ZM_0.6_1.0_KB_Zea_mays		genomic	clone ZMMBT480L01,					
VERSION	CC400575								
KEYWORDS	CC400575.1	GI:30880665							
SOURCE	GSS.								
ORGANISM	Zea mays								
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.								
AUTHORS	1 (baaes 1 to 797)								
	WhiteLaw,C.A., Quackenbush,J., Van Aken,S., Uteerback,T.,								
	Benner,A., Frazer,C.M., Yuan,Y., San Miguel,P., Ma,D. and								
	Reisenick,J.								
TITLE	Maize Genomics Consortium								
JOURNAL	Unpublished (2003)								
COMMENT	Other_GSSs: PURLHUTB								
	Contact: Cathy WhiteLaw								
	TIGR								
	9712 Medical Center Drive, Rockville, MD 20850, USA								
	Tel: 301-838-5843								
	Fax: 301-838-0208								
	Email: whiteLaw@tigr.org								
	Seq primer: TP								
	Class: sheared ends.								
FEATURES									
source	location/Qualifiers								
	1..797								
	/organism="Zea mays"								
	/mol_type="genomic DNA"								
	/strain="B73"								
	/db_xref="taxon:4577"								
	/clone="ZMMBT480L01"								
	/clone_1lb="ZM_0.6_1.0_KB"								
	/note="Vector: PCR4-TOPO, Site_1: EcoRI, 0.6-1.0 kb high								
	cot selected genomic DNA library"								
ORIGIN									
Query Match	5.2%	Score 86.6	DB 8	Length 797					
Best Local Similarity	59.7%	Pred. No. 8.2e-08							
Matches 163	Conservative 0	Mismatches 109	Indels 1	Gaps 1					
QY	713	TTTATAAATGACTATPAAGTCATTTTATATATAAGACACGCATGTCGTAGTTCTGCTT							772
DB	577	TTTAAAGACGCTTGSTAGACATGAGAAAGATATAGAAAAATATATCTTTTAAGAAATGTTA							518
QY	773	CAAAAATCTTCTGATTTTTTTTAAGAGCTAGTTTGCAACCCCTGTTCTTTCAAAGAATT							832
DB	517	TAAAGACAAAGAAACATTTTTTTTAAGAGCTAGTTTGTAATCCATTTTTTTCAAAAGATT							458
Q									

Db	457	TTCAATTCACAGAGGAAAAAATAGTTATTTTCCTTTGGTAAAAAGTAAATGACTCGGAA	398		
Qy	892	AAATGAGTTGCCAGACTAGCCCTTGAAATGTTTTCCCAATAAATTAACAATCACTGTGAT	951		
Db	397	AAATGAGTTCTCAAACTAGCCCTTAGGTTATATCGAATTAATCCCTATCTCATTTCTCT	338		
Qy	952	AATATTTGGCAGGCCCATTAATATTTAATC	984		
Db	337	ATTTCAAATTTCCACTTTGGCAAAATAGTGTCATC	305		
RESULT 11					
LOCUS	CC400574	820 bp	DNA linear GSS 19-MAY-2007		
DEFINITION	PNUH061TB ZM 0.6 1.0 KB Zea mays genomic clone ZM061A480L01,				
ACCESSION	CC400574	Genomic survey sequence.			
VERSION	CC400574.1	GI:30880664			
KEYWORDS	GSS.				
SOURCE	Zea mays				
ORGANISM	Zea mays				
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.					
1 (bases 1 to 820)					
Whitejak,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and Bennezen,J.					
Maize Genomics Consortium					
Unpublished (2003)					
Other_GSSs: PNUH061TD					
COMMENT					
Contact: Cathy Whitejak					
TIGR					
9712 Medical Center Drive, Rockville, MD 20850, USA					
Tel: 301-838-5843					
Fax: 301-838-0208					
Email: whitejak@tigr.org					
Seq primer: TR					
Class: sheared ends.					
FEATURES					
Location/Qualifiers					
1..820					
/organism="Zea mays"					
/mol_type="genomic DNA"					
/strain="B73"					
/db_xref="taxon:4577"					
/clone="ZM061A480L01"					
/clone_1lb="ZM 0.6 1.0 KB"					
/note="vector: pCR4-TOP0; Site 1: EcoRI; 0.6-1.0 kb high cot selected genomic DNA library"					
ORIGIN					
Query Match 6.2%; Score 86.6; DB 8; Length 820;					
Best Local Similarity 59.7%; Pred. No. 8,2e+08;					
Matches 163; Conservative 0; Mismatches 109; Indels 1; Gaps 1					
Qy	713	TTTATAAATGACTATAAAGTCATTTTATATTAAGACGCGATGCTGATCTCGTT	772		
Db	374	TTTAGAGGACGTTGAGATGAGAAAGATATAGAAATATATCTTTTAGAGAAATGTTA	433		
Qy	773	CAAAAATCTTCTGATTTTTTTAAGACTATTTGGCAACCTGTTCTTCCAAGANT	832		
Db	434	TAAAGGACAAACAATTTTTTAAGAGCTAGTTGGTATCCCATTTTTTCAAAAGATT	493		
Qy	833	TTGATTTTTTCAAAA-AAATTAAGTTTATTTTCTTTATATAAATAGAAAACACTTAGNA	891		
Db	494	TTCAATTCACAGAGGAAAAATAGTTTATTTCCCTTTGTAAAAATGTAATCACTCGGAA	553		
Qy	892	AAATGAGTTGCCAGACTAGCCCTTGAAATGTTTTCCCAATAAATTAACAATCACTGTGAT	951		
Db	554	AAATGAGTTCTCAAACTAGCCCTTAGGTTATATCGAATTAATCCCTATCTCATTTCTCT	613		
Qy	952	AATATTTGGCAGGCCCATTAATATTTAATC	984		

DB	614	ATTTCAAATTCACCTTGGCAAAATAGTGTCATC	646
RESULT 12	CG103452	471 bp	DNA linear GSS 20-AUG-2003
LOCUS	PUBB19TB.ZM.0.6.1.0_KB_Zea_mays	genomic clone ZMBBTA0628D13,	
DEFINITION	genomic survey sequence.		
ACCESSION	CG103452		
VERSION	CG103452.1	GI:33985746	
KEYWORDS	GSS.		
SOURCE	Zea mays		
ORGANISM	Zea mays		
REFERENCE	Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD clade; Panicoidae; Andropogoneae; Zea.		
AUTHORS	1 (bases 1 to 471) Whitelaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T., Reinick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and Benetzen,J.		
TITLE	Maize Genomics Consortium		
JOURNAL	Unpublished (2003)		
COMMENT	Other GSSs: PUBB19TD Contact: Cathy Whitelaw TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA Tel: 301-838-5843 Fax: 301-838-0208 Email: whitelaw@cigr.org Seq primer: 7R Clase: sheared ends. Location/Qualifiers 1..471 /organism="Zea mays" /mol_type="genomic DNA" /strain="873" /db_xref="taxon:4577" /clone="ZMBBTA0628D13" /clone_11b="ZM.0.6.1.0_KB" /note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high cot selected genomic DNA library"		
FEATURES			
SOURCE			
ORIGIN			
Query Match	6.1%	Score 85.4;	DB 9; Length 471;
Best Local Similarity	69.5%;	Pred. No. 1.4e-07;	
Matches 116;	Conservative 0;	Mismatches 51;	Indels 0; Gaps 0;
QY	789	TTTTTAAAGACTAGTTGGCAACCCCTGTTCTTTCAGAAATTTGATTTTTCAAAA	848
DB	216	TTTTCTTAGACTAGTTGGAAATCCATTTTTTTCAGGATTTACATTTTTTTCAGGG	275
QY	849	AAATAGTTTATTTTCTCTTATATAAATAGAAACACTTAGAAAAATAGAGTGCAGAC	908
DB	276	AAATAGTTCAATTTCTTGTAAAAAATAGAAATACATGGGAAATATGTTTCCCAAC	335
QY	909	TAGCCCTAGAAATGTTTCCCAATAAATTACATCATCTGTATTAAT	955
DB	336	TAGCGCTTAGACTATTTTCTAAAAAAGTAGATCCAAATGATTAAGATT	382
RESULT 13	CG082135	765 bp	DNA linear GSS 20-AUG-2003
LOCUS	PUBO12TD.ZM.0.6.1.0_KB_Zea_mays	genomic clone ZMBBTA0712B24,	
DEFINITION	genomic survey sequence.		
ACCESSION	CG082135		
VERSION	CG082135.1	GI:33964429	
KEYWORDS	GSS.		
SOURCE	Zea mays		
ORGANISM	Zea mays		
REFERENCE	Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD clade; Panicoidae; Andropogoneae; Zea.		

```

REFERENCE      TITLE              JOURNAL COMMENT
AUTHORS        WhiteLaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T., Reenick,A., Frazer,C.M., Yuan,Y., San Miguel,P., Ma,J. and Benmetzen,J.
Maize Genomics Consortium
Unpublished (2003)
Other GSSS: PUFQX12TB
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@cigr.org
Seq primer: TP
Class: sheared ends.

FEATURES
source          location/Qualifiers
                +-----+
                |       .765       |
                | /organism="Zea mays"|
                | /mol_type="genomic DNA"|
                | /strain="B73"         |
                | /db_xref="taxon:4577"|
                | /clone="ZMMBT0712B24" |
                | /clone_lib="ZM_0.6_1.0 KB" |
                | /note="Vector:pCR4-TORO; Site_1:EcoRI; 0.6-1.0 kb high|
                | cor selected genomic DNA library"   |
                +-----+

ORIGIN
Query Match    6.1%; Score 85.2; DB 9; Length 765;
Best Local Similarity 71.8%; Pred. No.1.6e-07;
Matches 125; Conservative 0; Mismatches 48; Indels 1; Gaps 1;

Dy     787 ATTTTAAAGACGCTAATTGGCAACCCTGTTCCTTAACAAGAATTTGAATTTTTCAA 846
            |||||
Db     255 ATTCTTGCAATGACTAGATTGGGAACCTTAATTT-TCTCAGCGGATTTTATAATTTTCCAA 197

Dy     847 AAAAATTAGTTTATTTCTCTTTATAAATAGAAAAACATTAGAAAATAGAGTTGCCAG 906
            |||||
Db     196 GAATAATAGTTTATTTCTCTGTGAAAAATAAATCATCTTGAGAAAAATATAGTTTCCA 137

Dy     907 ACTAGCCCCGAATAGTTTCCCANTAAATTAATCATCACTGTGTATAATTTATTTT 960
            |||||
Db     136 ACTAGCCCTTAAGATTTTATATAATATGAGGAAAAAATTCCTCTATTTTTCATTG 83

RESULT 14
CC630219/c
LOCUS      CC630219               781 bp           DNA             linear      GSS 19-JUN-2003
DEFINITION OSUCGSJTV_ZM_0.7_1.5_KB Zea mays genomic clone ZMBRHd0406U09,
VERSION    CC630219
KEYWORDS   genomic survey sequence.
SOURCE     CC630219.1 GI:32003354
ORGANISM   Zea mays
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 781)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T., Reenick,A., Frazer,C.M., Budiman,M.A., Bedell,J.R., Kohlfing,T., Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSS: OSUCGSJTH
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@cigr.org
Seq primer: TP
Class: sheared ends.
Location/Qualifiers
```

Search completed: September 16, 2005, 08:08:41
Job time : 4348.69 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using bw model

Run on: September 15, 2005, 08:25:31 ; Search time 526.823 Seconds

(without alignment)
1440.280 Million cell updates/sec

Title: US-10-713-381-2_COPY_1155_1311

Perfect score: 157
Sequence: 1 cgcgcacatccacatgcacat.....cgcacatattcttcctgaac 157

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hgt:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sbs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	157	100.0	255	6	AX2244402 Sequence
2	157	100.0	1394	6	AX224394 Sequence
3	157	100.0	1394	6	AX224395 Sequence
4	157	100.0	1394	6	BD062176 Male t1s
5	157	100.0	1394	6	BD062177 Male t1s
6	157	100.0	3343	8	AF360356 Zea mays
7	146	93.0	158	6	AX224396 Sequence
8	54	34.4	137327	8	AC135206 Oryza sat
9	50	31.8	50	6	AX224398 Sequence
10	40	25.5	40	6	AX224399 Sequence
11	34.8	22.2	174806	2	BX950854 Danio rer
12	34.6	22.0	110127	9	AC104134 Homo sapi
13	34.4	21.9	214946	2	CR847532 Homo sapi
14	33.8	21.5	141810	2	CR847533 Danio rer
15	33.8	21.5	214178	5	BX640577 Zebrafish
16	33.2	21.1	117764	9	AC011290 Homo sapi
17	33.2	21.1	162481	2	AC091864 Homo sapi
18	33.2	21.1	182069	2	AC146134 Pan trogl
19	32.8	20.9	82642	9	AC103745 Homo sapi

C	20	32.8	20.9	93610	9	AC006602 Homo sapi
C	21	32.8	20.9	130031	2	AC135264 Homo sapi
C	22	32.8	20.9	148598	9	HSBA51C14 Human DNA
C	23	32.8	20.9	158316	9	AC135350 Homo sapi
C	24	32.8	20.9	169771	9	AC087485 Homo sapi
C	25	32.8	20.9	169667	2	AC019229 Homo sapi
C	26	32.8	20.9	193569	2	AC149044 Pan trogl
C	27	32.6	20.8	53688	2	AC099973 Mus muscu
C	28	32.6	20.8	146080	5	AL953855 Zebrafish
C	29	32.6	20.8	165845	2	AC127012 Rattus no
C	30	32.6	20.8	186629	2	AC147308 Pan trogl
C	31	32.6	20.8	204190	2	AC140696 Rattus no
C	32	32.6	20.8	252534	2	AC097405 Rattus no
C	33	32.4	20.6	2070	10	BC020027 Mus muscu
C	34	32.4	20.6	2096	10	BC038066 Mus muscu
C	35	32.4	20.6	2433	10	BC055469 Mus muscu
C	36	32.4	20.6	110000	1	AB017197_05 Continuation (6 of
C	37	32.4	20.6	137832	2	AC104550 Mus muscu
C	38	32.4	20.6	168889	10	AL731779 Mouse DNA
C	39	32.4	20.6	182847	2	AC148229 Gallus ga
C	40	32.4	20.6	221338	5	AC147437 Gallus ga
C	41	32.4	20.6	226440	2	AC139057 Mus muscu
C	42	32.2	20.5	139173	9	AC009140 Homo sapi
C	43	32.2	20.5	150116	2	AC034132 Homo sapi
C	44	32.2	20.5	189323	2	AC150620 Callithrix
C	45	32.2	20.5	205954	9	AC008546 Homo sapi

ALIGNMENTS

RESULT 1	AX224402	255 bp	DNA	linear	PAT 10-SEP-2001
LOCUS	AX224402	Sequence 9 from Patent WO0160997.			
DEFINITION	AX224402				
ACCESSION	AX224402				
VERSION	AX224402.1	GI:15554644			
KEYWORDS					
SOURCE					
ORGANISM					
	Zea mays				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.				
REFERENCE					
AUTHORS	Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.				
TITLE	Male tissue-preferred regulatory region and method of using same				
JOURNAL	Patent: WO 0160997-A 9 23-AUG-2001;				
	PIONEER HI-BRED INTERNATIONAL, INC. (US)				
FEATURES					
source					
	1..255				
	/organism="Zea mays"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:4577"				
ORIGIN					
Query Match	100.0%; Score 157; DB 6; Length 255;				
Best Local Similarity	100.0%; Pred.No. 1.8e-40;				
Matches	157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	1	CGTGTTCATTCACATGCACTACTACATGCTTGTTCACCGCTTCGCTTGTTCATTCGTC	60		
DB	15	CGTGTTCATTCACATGCACTACTACATGCTTGTTCACCGCTTCGCTTGTTCATTCGTC	74		
QY	61	AAGCCTTGCTATTCCTGAACCAAGAGATCACTACTCCCAACATTCATTCATCATG	120		
DB	75	AAGCCTTGCTATTCCTGAACCAAGAGATCACTACTCCCAACATTCATTCATCATG	134		
QY	121	CAACTTCATGCAACACGACATATGTTCTGTAAC	157		
DB	135	CAACTTCATGCAACACGACATATGTTCTGTAAC	171		
RESULT 2					

AX224394
LOCUS AX224394 1394 bp DNA linear PAT 10-SEP-2001
DEFINITION Sequence 1 from Patent WO0160997.
ACCESSION AX224394
VERSION AX224394.1 GI:15554636
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
Clade; Panicoidae; Andropogoneae; Zea.
REFERENCE
1 Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
Patent: WO 0160997-A 1 23-AUG-2001.
PIONEER HI-BRED INTERNATIONAL, INC. (US)
JOURNAL
FEATURES
source Location/Qualifiers
1..1394
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"
ORIGIN
Query Match 100.0%; Score 157; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 1.9e-40;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGTGCATCTCAGATGGCATCTACTACACTGCTGTTCAACCGTTGCTTTCATCGTCC 60
DB 1155 CGTGCATCTCAGATGGCATCTACTACACTGCTGTTCAACCGTTGCTTTCATCGTCC 1214
QY 61 AAGCCTTGCTATTCTGAAACCAAGAGATACCTACTCCCAACAATCCATCTTACTCAG 120
DB 1215 AAGCCTTGCTATTCTGAAACCAAGAGATACCTACTCCCAACAATCCATCTTACTCAG 1274
QY 121 CACTTCATGCAACACGACATATGTTCTGAAAC 157
DB 1275 CACTTCATGCAACACGACATATGTTCTGAAAC 1311
RESULT 3
AX224395 1394 bp DNA linear PAT 10-SEP-2001
LOCUS AX224395
DEFINITION Sequence 2 from Patent WO0160997.
ACCESSION AX224395
VERSION AX224395.1 GI:15554637
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
Clade; Panicoidae; Andropogoneae; Zea.
REFERENCE
1 Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
Patent: WO 0160997-A 2 23-AUG-2001.
PIONEER HI-BRED INTERNATIONAL, INC. (US)
JOURNAL
FEATURES
source Location/Qualifiers
1..1394
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"
ORIGIN
Query Match 100.0%; Score 157; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 1.9e-40;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGTGCATCTCAGATGGCATCTACTACACTGCTGTTCAACCGTTGCTTTCATCGTCC 60
DB 1155 CGTGCATCTCAGATGGCATCTACTACACTGCTGTTCAACCGTTGCTTTCATCGTCC 1214
QY 61 AAGCCTTGCTATTCTGAAACCAAGAGATACCTACTCCCAACAATCCATCTTACTCAG 120

DB 1215 AAGCCTTGCTATTCTGAAACCAAGAGATACCTACTCCCAACAATCCATCTTACTCAG 1274
QY 121 CACTTCATGCAACACGACATATGTTCTGAAAC 157
DB 1275 CACTTCATGCAACACGACATATGTTCTGAAAC 1311
RESULT 4
BD062176 1394 bp DNA linear PAT 27-AUG-2002
LOCUS BD062176
DEFINITION Male tissue-preferred regulatory region and method of using same.
ACCESSION BD062176
VERSION BD062176.1 GI:22607781
KEYWORDS JP 2001520523-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1394)
Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G.A. and Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
Patent: JP 2001520523-A 1 30-OCT-2001.
PIONEER HI BRED INTERNATIONAL INC
PN JP 2001520523-A/1
PD 30-OCT-2001
PF 19-JUN-1998 JP 1999504910
PR 23-JUN-1997 US 08/880499
PI MARC C ALBERTSEN,TIMOTHY W FOX,CARL W GARNAAT,GARY A HUFFMAN,
PI TIMMY L KENDALL
PC C12N15/62,C12N15/29,C12N9/24,C12N9/22,C12N9/10,C12N9/00 PC
PC C07K14/34,C12Q1/68,
PC A01H5/00
CC Strandedness: Single;
CC Topology: linear;
FH key Location/Qualifiers
1..1394
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 157; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 1.9e-40;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGTGCATCTCAGATGGCATCTACTACACTGCTGTTCAACCGTTGCTTTCATCGTCC 60
DB 1155 CGTGCATCTCAGATGGCATCTACTACACTGCTGTTCAACCGTTGCTTTCATCGTCC 1214
QY 61 AAGCCTTGCTATTCTGAAACCAAGAGATACCTACTCCCAACAATCCATCTTACTCAG 120
DB 1215 AAGCCTTGCTATTCTGAAACCAAGAGATACCTACTCCCAACAATCCATCTTACTCAG 1274
QY 121 CACTTCATGCAACACGACATATGTTCTGAAAC 157
DB 1275 CACTTCATGCAACACGACATATGTTCTGAAAC 1311
RESULT 5
BD062177 1394 bp DNA linear PAT 27-AUG-2002
LOCUS BD062177
DEFINITION Male tissue-preferred regulatory region and method of using same.
ACCESSION BD062177
VERSION BD062177.1 GI:22607782
KEYWORDS JP 2001520523-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1394)

AUTHORS	TITLE	COMMENT	FEATURES	ORIGIN
ALBERTSEN, M. C., FOX, T. W., GARMAAT, C. W., HUFFMAN, G. A. and KENDALL, T. L.	Male tissue-preferred regulatory region and method of using same patent: JP 2001530523-A 2 30-OCT-2001;			
PIONEER HI BRED INTERNATIONAL INC				
PN JP 2001530523-A/2				
PD 30-OCT-2001				
PF 19-JUN-1998 JP 1999504910				
PR 23-JUN-1997 US 08/860499				
PI MARC C ALBERTSEN, TIMOTHY W FOX, CARL W GARMAAT, GARY A HUFFMAN, PI TIMMY L KENDALL				
PC C12N15/22, C12N15/29, C12N9/24, C12N9/22, C12N9/10, C12N9/00 PC, C07K14/34, C12Q1/68,				
PC A01H5/00				
CC Strandedness: Single;				
CC Topology: Linear;				
FH Key				
Location/Qualifiers				
1..1394				
/organism="Homo sapiens"				
/mol_type="Genomic DNA"				
/db_xref="taxon:9606"				
Query Match	100.0%; Score 157; DB 6; Length 1394;			
Best Local Similarity	100.0%; Pred. No. 1.9e-40;			
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
OY 1 CGTGCATCTCAGCAGGATCTACTATCTGTTTCACCGTTCGCTGTTTCATCTGCC 60				
DB 1155 CGTGCATCTCAGCAGGATCTACTATCTGTTTCACCGTTCGCTGTTTCATCTGCC 1214				
OY 61 AAGCGTTCGCTATTTGGAACAGAGATACCTACTCCCAACATCATCTTACTCATG 120				
DB 1215 AAGCGTTCGCTATTTGGAACAGAGATACCTACTCCCAACATCATCTTACTCATG 1274				
OY 121 CACTTCATGCAACAGCAGCATATGTGTTCTGTGAC 157				
DB 1275 CACTTCATGCAACAGCAGCATATGTGTTCTGTGAC 1311				
RESULT 6				
AF360356	3343 bp DNA linear	PLN 12-MAY-2001		
LOCUS				
DEFINITION	Zea mays male fertility protein (M645) gene, complete cds.			
ACCESSION	AF360356			
VERSION	AF360356.1 GI:14028756			
KEYWORDS				
SOURCE				
ORGANISM				
	Zea mays			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoidae; Andropogoneae; Zea.			
	1 (bases 1 to 3343)			
	Fox, T. W., Trimmell, M. R. and Albertsen, M. C.			
	Cloning of M645, a gene required for male fertility from Zea mays unpublished			
	2 (bases 1 to 3343)			
	Fox, T. W., Trimmell, M. R. and Albertsen, M. C.			
	Direct Submission			
	Submitted (13-MAR-2001) Trait and Technology Development, Pioneer Hi-Bred Intl. Inc., 7300 N.W. 62nd Ave., P.O. Box 1004, Johnston, IA 50131-1004, USA			
	1A 50131-1004, USA			
FEATURES				
source				
	1..3343			
	/organism="Zea mays"			
	/mol_type="Genomic DNA"			
	/cultivar="B73"			
	/db_xref="taxon:4577"			
	/chromosome="9L"			
	<1392..>2842			
	gene			
	join(<13392..1768,1898..2182,2280..2447,2534..>2942)			

CDs

/gene="W645"
/product="male fertility protein"
join(1392..1168,1898..2182,2280..2447,2534..2942)
/gene="W645"
/codon_start=1
/product="male fertility protein"
/protein_id="AAK52489.1"
/db_xref="GI:14028757"
/translation="MEKRNILQWRGRDGIYQVPHLFFAALALALLVADPEGLSPLAEV
DYRVKHEIAPYGEVMSWPNLDNRSLRGRLEGVGVPEBESI EPLDQKGPYAGLAE
DGRVVRMGEAPGEWETFEAVNVPDMSREVCAQNVSTTRKHQEKHEPCGRLGFRHEE
TGEIYVADATYGLMVTQSGSVASSVAREADGDPIRANLIDIRNKSVPFTDTSNRY
SRKHNLILLEGEGRLRLRPDPTSGVHVLKGLVFPNGVQVISEDQFLFSETTNC
RIMYKVEGPAGEVEVEFANLPFPDMVRNSRGQFWVAIDCCTPDAOEYFAKRPMLR
TLVYKFPPLSLKVLITWKAARMHVTVALIDEGRVAEVLEDRGHEVMKLVSREVGRKK
LMIGTVAHNHIATIPLEPD"

ORIGIN

Query Match 100.0%; Score 157; DB 8; Length 3343;
Best Local Similarity 100.0%; Pred. No. 1.9e-40;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGTGTCAATTCACATGGCATCTACATCAAGCTTTGTTCACACCGTTGCTCTTTGTTCATCGTCC 60
DB 1155 CGTGTCAATTCACATGGCATCTACATCAAGCTTTGTTCACACCGTTGCTCTTTGTTCATCGTCC 1214
OY 61 AAGCTTGCCATATCTGAAACCAAGAGATACCTACTCCCAACCAATCATCTTACTCATG 120
DB 1215 AAGCTTGCCATATCTGAAACCAAGAGATACCTACTCCCAACCAATCATCTTACTCATG 1274
OY 121 CAACCTTCATGCAAAACGACCATATGTCTTCTGAAC 157
DB 1275 CAACCTTCATGCAAAACGACCATATGTCTTCTGAAC 1311

RESULT 7
AX224396 158 bp DNA linear PAT 10-SEP-2001
LOCUS Sequence 3 from Patent WO0160997.
DEFINITION AX224396
ACCESSION AX224396.1 GI:15554638
KEYWORDS
XREFS
SOURCE
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
Clade; Panicoideae; Andropogoneae; Zea.
1
REFERENCE
AUTHORS Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
TITLE Male tissue-preferred regulatory region and method of using same
JOURNAL Patent: WO 0160997-A 3 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
source
1..158
Location/Organisms
/organism="Zea mays"
/mol_type="unasigned DNA"
/db_xref="taxon:4577"

ORIGIN

Query Match 93.0%; Score 146; DB 6; Length 158;
Best Local Similarity 99.4%; Pred. No. 7.2e-37;
Matches 157; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 1 CGTGTCAATTCACATGGCATCTACATCAAGCTTTGTTCACACCGTTGCTCTTTGTTCATCGTCC 59
DB 1 CGTGTCAATTCACATGGCATCTACATCAAGCTTTGTTCACACCGTTGCTCTTTGTTCATCGTCC 60
OY 60 CAAGCTTGCCATATCTGAAACCAAGAGATACCTACTCCCAACCAATCATCTTACTCAT 119
DB 61 CAAGCTTGCCATATCTGAAACCAAGAGATACCTACTCCCAACCAATCATCTTACTCAT 120
OY 120 GCAACTTCATGCAAAACGACCATATGTCTTCTGAAC 157

Db	121	GCAACTTCATGCAACAGCACATATGTTTCTGAAC	158
RESULT 8	AC135206	137327 bp	DNA
LOCUS	Oryza sativa (japonica cultivar-group)	1linear	PLN 16-APR-2003
DEFINITION	OJ1041F02, complete sequence.	chromosome 3 clone	
ACCESSION	AC135206		
VERSION	AC135206.3	GI:27596977	
SOURCE	HTG.		
ORGANISM	Oryza sativa (japonica cultivar-group)		
REFERENCE	Oryza sativa (japonica cultivar-group)		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
REFERENCE	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;		
REFERENCE	Ehrhartoideae; Oryzaceae; Oryza.		
REFERENCE	1 (bases 1 to 137327)		
REFERENCE	Wing,R.A., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Currie,J. and		
REFERENCE	Collura,K.		
REFERENCE	Rice Genomic Sequence		
REFERENCE	2 (bases 1 to 137327)		
REFERENCE	Wing,R.A., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Saeki,C.,		
REFERENCE	Currie,J., Collura,K. and Thompson,S.		
REFERENCE	Direct Submision		
REFERENCE	Submitted (09-OCT-2002) Arizona Genomics Institute, University of		
REFERENCE	Arizona, 303 Forbes, Tucson, AZ 85721, USA		
REFERENCE	3 (bases 1 to 137327)		
REFERENCE	Wing,R.A., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Saeki,C.,		
REFERENCE	Currie,J., Collura,K. and Thompson,S.		
REFERENCE	Direct Submision		
REFERENCE	Submitted (06-NOV-2002) Arizona Genomics Institute, University of		
REFERENCE	Arizona, 303 Forbes, Tucson, AZ 85721, USA		
REFERENCE	4 (bases 1 to 137327)		
REFERENCE	Wing,R.A., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Currie,J.,		
REFERENCE	Collura,K. and Thompson,S.		
REFERENCE	Direct Submision		
REFERENCE	Submitted (11-JAN-2003) Arizona Genomics Institute, University of		
REFERENCE	Arizona, 303 Forbes, Tucson, AZ 85721, USA		
REFERENCE	5 (bases 1 to 137327)		
REFERENCE	Wing,R.A., Yu,Y., Soderlund,C., Chen,M., Kim,H., Rambo,T.,		
REFERENCE	Saeki,C., Henry,D., Thompson,S., Simmons,J., Thurmond,S.K. and		
REFERENCE	Sun,S.		
REFERENCE	Direct Submision		
REFERENCE	Submitted (16-APR-2003) Clemson University Genomics Institute,		
REFERENCE	Clemson University, 100 Jordan Hall, Clemson, SC 29634, USA		
REFERENCE	On Jan 11, 2003 this sequence version replaced gi:24635891.		
REFERENCE	This sequence was finished as follows unless otherwise noted: all		
REFERENCE	regions were either double-stranded or sequenced with an alternate		
REFERENCE	chemistry or covered by high quality data (i.e., phred quality		
REFERENCE	>30); an attempt was made to resolve all sequencing problems, such		
REFERENCE	as compressions and repeats; all regions were covered by more than		
REFERENCE	one plasmid subclone; and the assembly was confirmed by		
REFERENCE	restriction digest. There are transposons located at area		
REFERENCE	2947-32874 and at area 19945-33212. Region 70511-70536 is covered		
REFERENCE	by a single subclone. Region 18678-18873 is a double stranded		
REFERENCE	single subclone. Areas 69957-70761, 89836-90416, 98951-99009,		
REFERENCE	105316-105607, and 133052-133152 are covered by Monsanto only. The		
REFERENCE	nucleotide sequence of this BAC clone was generated by combining		
REFERENCE	SynGene, Monsanto and Arizona Genomics Institute sequencing data.		
REFERENCE	Location/Qualifiers		
REFERENCE	1. 137327		
REFERENCE	/organism="Oryza sativa (japonica cultivar-group)"		
REFERENCE	/mol_type="genomic DNA"		
REFERENCE	/db_xref="taxon:39947"		
REFERENCE	/chromosome="3"		
REFERENCE	/clone="OJ1041F02"		
REFERENCE	/note="(japonica cultivar-group)"		
REFERENCE	/complement(638..9495)		
REFERENCE	/gene="OJ1041F02.1"		
REFERENCE	/note="unknown protein"		
REFERENCE	/complement(Join(638..838,1141..1276,3133..3244,3326..3421,		
REFERENCE	3513..3727,3812..3934,4024..4054,4157..4216,5088..5243,		
REFERENCE	5322..5411,5484..5699,5808..5942,6035..6538,6622..6837,		
REFERENCE	7548..7593,7679..7815,8196..8328,9221..9495))		
REFERENCE	/gene="OJ1041F02.1"		
REFERENCE	/codon_start=1		
REFERENCE	/product="unknown protein"		
REFERENCE	/protein_id="AAP06844.1"		
REFERENCE	/db_xref="GI:29893590"		
REFERENCE	/translation="MAOSLELLIOFLMPDNDARROAEQIRRLARDPPVVALVHNL		
REFERENCE	RTAKTPNRQLAAVILRRKITSHPKLEPPHAKASIKQALLDSITTDHSHLVARSANV		
REFERENCE	VSIINKAVPAGMPELLEPFIQCSQSFQEDHREVALILFSLTETIGTFOSHLDAN		
REFERENCE	OPTILKCLQDEKSSRRVIALKAVGSFLIEVNDGDVVKIFEDPVSILTVNSQCLAN		
REFERENCE	GEEVDASIAPELDELIESPAVPLDSDVSRYQFLEVCNSDELINIRQQAIIQITIS		
REFERENCE	LVKFASFLKKGKLVIPILQVCPILTTETADDEGSDLAADBSAEITDVMALNIPRH		
REFERENCE	VPPPLAEASVSPRINPKYEAATATSGVSEGCCHKDLDECLKVLALADDE		
REFERENCE	OMVRASAPAGQFAPEHLPEPLISVSESVLPCTINALDPSDEVEKSYVALAACED		
REFERENCE	MGENILPLDPLMKRLVNSLOGSPNNLETCSAIGSVAALAEQAFMYAEVLEMK		
REFERENCE	GFMYITNDEDLCARARATEVVGIVMAAGRAMETILPEPFLAALISGVLDSELEY		
REFERENCE	THGFPSVNAELIDDSFAQYLPHVPLAASGCLDSDGSAVDIDDADVNSGFGVSDD		
REFERENCE	DVNDPRVRNISMLTIIADIILTAIPAPADVLEKQDILDTVMNTIYIKTRRDD		
REFERENCE	KEVVAOACTSLADIVRDCGPAIIIEYIRIADATLILROSCCOQVESDDEGDDID		
REFERENCE	HDEVIMDAVSPDLPAFAVMGSPDPPIFTKLFDLSMKRAKSPHPQDCTMYATLAEV		
REFERENCE	AQMGAPISATYDKIMPLVKEIASSEATRRNRNAFCYEMCKNGAALAKYIGVILH		
REFERENCE	GLHRLFADESEPDADVADNAAGAIAMINVQQSIPLNQIILPLVPVINAFAQVVSFN		
REFERENCE	ESDEKVTVAKAVSHLISVYGQMPILSALPFAHNAALASFANRR"		
REFERENCE	12372..15503		
REFERENCE	/note="unknown protein"		
REFERENCE	/join(12372..12757,12858..13025,13130..13295,13716..13826,		
REFERENCE	13982..14130,14224..14354,14459..14558,15197..15503)		
REFERENCE	/gene="OJ1041F02.2"		
REFERENCE	/codon_start=1		
REFERENCE	/product="unknown protein"		
REFERENCE	/protein_id="AAP06845.1"		
REFERENCE	/db_xref="GI:29893591"		
REFERENCE	/translation="MPPOGSECRRHGWLGVEBAVGGGEPPVPLRKRSVDKAS		
REFERENCE	TAPRICWECDEAGDITCDIVAPLRSCSAKAMPPEAPLPFRMTPEPPRQRDGE		
REFERENCE	BARPGEAIIKRGKHSYSLNLQIGISVSYSKSHLPQKLAASDPDEPREKWTFFP		
REFERENCE	BGSKTTPHNSVDPEWKDQCPAVPHRKLKLYEVDYEAVMIAICGDTLRBLASKCP		
REFERENCE	SCFRTQDPRMIVTKKSEYKVLIRMSRYEVAROVKSTLLRPFTCHTCKQKCP		
REFERENCE	KIHRRFDLKGSHGRTIDTKRKIDETTLKDLQDQVAFRIQRFYEELMKQIQDCT		
REFERENCE	FLETQGLMDYSLGLVHFRNDYVSQKIGISQHLAPKSTGRKSKSEGSFCELCFVE		
REFERENCE	SGCKDRDLIDBRKEPITQGIINMPQAEKSSKILDNPLNHRHLITPSPSGSCDYVL		
REFERENCE	FFGIIDLDQDYDTKLEHAYKSFQVNPDIYSAVDPKLISRFPQFIRFVIFKEQ"		
REFERENCE	16734..16878		
REFERENCE	/note="putative MITE, Gaijin/Gaigin-like"		
REFERENCE	/join(19263..19454		
REFERENCE	/note="putative transposon, CACTA, En/Spm-like"		
REFERENCE	21819..24660		
REFERENCE	/gene="OJ1041F02.3"		
REFERENCE	/note="hypothetical protein"		
REFERENCE	/join(21819..22007,22096..22206,23772..23996,24523..24660)		
REFERENCE	/gene="OJ1041F02.3"		
REFERENCE	/codon_start=1		
REFERENCE	/product="hypothetical protein"		
REFERENCE	/protein_id="AAP06846.1"		
REFERENCE	/db_xref="GI:29893592"		
REFERENCE	/translation="MATASDGGSGGGEORRLLSIPKEGERIIPATRRPDGTRKAIIRI		
REFERENCE	PAGYPROEVAIYQSKAQKSGSDVPCYDPAIDAPKTKAAKRNREKREKQAS		
REFERENCE	TTNDYGKGLHIEDAGETDNKDAVDSYTKQISGIAISESLVATSYTDADNSES		
REFERENCE	SAPIDDKIRALKKIRLAEAOVGDPENLKPEOLEKWKIKEGWKEELKLENKSSPA		
REFERENCE	AS"		
REFERENCE	/complement(26563..30575)		
REFERENCE	/gene="OJ1041F02.4"		
REFERENCE	/note="unknown protein"		
REFERENCE	/complement(Join(26563..26772,27082..27288,27430..27537,		
REFERENCE	27881..27940,28051..28126,28327..28484,28616..28678,		
REFERENCE	28790..28874,29560..30575))		
REFERENCE	/gene="OJ1041F02.4"		
REFERENCE	/codon_start=1		
REFERENCE	/product="unknown protein"		
REFERENCE	/protein_id="AAP06847.1"		
REFERENCE	/db_xref="GI:29893593"		
REFERENCE	/translation="MGPPAAPDPSSSSSGSSRRRRRLRLRRMAAKHIGYDASNFCAYP		

KEYWORDS	HTG: HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
SOURCE	Danio rerio (zebrafish)
ORGANISM	Danio rerio
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
REFERENCE	1 (bases 1 to 174806)
AUTHORS	McLaren, S.
TITLE	Direct Submission
JOURNAL	Submitted (28-SEP-2004) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfish-help@sanger.ac.uk
COMMENT	zfish-help@sanger.ac.uk On Sep 25, 2004 this sequence version replaced gi:42821090.

all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPc1-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P. Y., Zhao, B., Frengen, E., Tatenio, M., Caranese, J. J. and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.regen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-554H10, 2000 bp overlap; the clone sequenced to the right is RP11-450B9, 2000 bp overlap. Actual start of this clone is at base position 151996 of RP11-554H10.

FEATURES

source

location/Qualifiers
1..110127
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RP11-525L16"
/clone_id="RPc1-11"
701..998
/rpt_family="Alu"
999..1297
/rpt_family="Alu"
1410..1486
/rpt_family="L1"
1564..1678
/rpt_family="Alu"
1684..1722
/rpt_family="AT_rich"
2338..2456
/rpt_family="MIR"
4039..4332
/rpt_family="Alu"
4412..4532
/rpt_family="Alu"
4669..4846
/rpt_family="MIR"
5878..6173
/rpt_family="Alu"
6385..6584
/rpt_family="MIR"
6929..7121
/rpt_family="MIR"
7167..7312
/rpt_family="MIR"
7953..8202
/rpt_family="Alu"
8537..8577
/rpt_family="L2"
8585..8756
/rpt_family="L1"
8768..8888
/rpt_family="Alu"
9005..9219
/rpt_family="L2"

repeat_region 9272..9838
/rpt_family="L1"
repeat_region 9849..9976
/rpt_family="Alu"
repeat_region 10124..10411
/rpt_family="Alu"
repeat_region 10483..10759
/rpt_family="L2"
repeat_region 11215..11264
/rpt_family="(TA)n"
repeat_region 11362..11476
/rpt_family="Alu"
repeat_region 11478..11776
/rpt_family="Alu"
repeat_region 12314..12607
/rpt_family="Alu"
repeat_region 12778..12943
/rpt_family="MALR"
repeat_region 12954..12979
/rpt_family="Alu"
repeat_region 14391..14655
/rpt_family="Alu"
repeat_region 14670..14939
/rpt_family="MALR"
repeat_region 15435..15449
/rpt_family="L2"
repeat_region 15450..15730
/rpt_family="Alu"
repeat_region 15731..15824
/rpt_family="L2"
repeat_region 15837..15915
/rpt_family="MALR"
repeat_region 16150..16214
/rpt_family="MALR"
repeat_region 16224..16382
/rpt_family="MER1_type"
repeat_region 16533..16747
/rpt_family="MIR"
repeat_region 16889..17116
/rpt_family="L2"
repeat_region 17271..17343
/rpt_family="L1"
repeat_region 17384..17534
/rpt_family="MIR"
repeat_region 17545..17843
/rpt_family="Alu"
repeat_region 17995..18056
/rpt_family="L2"
repeat_region 18057..18074
/rpt_family="GA_rich"
repeat_region 18075..18379
/rpt_family="Alu"
repeat_region 18380..18404
/rpt_family="GA_rich"
repeat_region 18889..19067
/rpt_family="MER1_type"
misc_feature 19296..19302
/note="match to EST BE843354 (NID:G10275732)"
repeat_region 19593..19709
/rpt_family="Alu"
repeat_region 19890..20158
/rpt_family="Alu"
repeat_region 20225..20698
/rpt_family="L1"
repeat_region 20737..20769
/rpt_family="AT_rich"
repeat_region 20887..21186
/rpt_family="Alu"
misc_feature 21350..21369
/note="match to EST BG571642 (NID:G13579295)"
repeat_region 21357..21652
/rpt_family="Alu"
repeat_region 21653..21957


```

repeat_region      22328..22636
                    /rpc_family="Alu"
Query Match       22.0%; Score 34.6; DB 9; Length 110127;
Best Local Similarity 54.3%; Pred. No. 3.1;
Matches 70; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 3 TGTGATCTGACATGAGGATCTACATGCTGTGTTACACGGTTCGTCCTGTTCCATCTGTCGCA 62
DB 8763 TGGGATTAACAAGCGTGAGCCAGTACCTGGCCAACTATCTTTTCTTCCTCTCTCCCA 87692
QY 63 GCGTTGCTTATTTGAAACCAAGAGGATCTACTATCCCAACATCATCTTACTATGCA 122
DB 87693 TCTCCCTTATTTTACCTCTTGTTCATCTACTCAGAACTCACCATTCTCCATACC 87752
QY 123 ACTTCGATG 131
DB 87753 ACCTCCAG 87761

RESULT 13
AC136896      214946 bp      DNA      linear      PRI 21-MAR-2003
LOCUS      Homo sapiens chromosome 15, clone RP11-86001, complete sequence.
DEFINITION      AC136896
VERSION      AC136896.6 GI:29135656
KEYWORDS      HTG.
SOURCE      Homo sapiens (human)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      Birren, B., Nusbaum, C. and Lander, E.
              1 (bases 1 to 214946)
              Homo sapiens chromosome 15, clone RP11-86001
              Unpublished
              2 (bases 1 to 214946)
              Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
              Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B.,
              Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
              Cook, A., Cooke, P., DeRellano, K., Dewar, K., Diaz, J. S., Dodge, S.,
              Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,
              Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hafez, N.,
              Hago, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
              Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
              Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J.,
              Matchew, C., McCarthy, M., Meldrim, J., Menus, L., Mihova, T.,
              Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
              Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,
              Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R.,
              Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupack, R.,
              Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
              Stojanovic, N., Talamas, J., Testaye, S., Theodore, J., Topham, K.,
              Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X.,
              Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
              Direct Submission
              Submitted (09-NOV-2002) Whitehead Institute/MIT Center for Genome
              Research, 320 Charles Street, Cambridge, MA 02141, USA
              3 (bases 1 to 214946)
              Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
              Anderson, S., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T.,
              Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y.,
              Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K.,
              Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
              Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S.,
              Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hago, B.,
              Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
              Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
              Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C.,
              Macdonald, P., Major, J., Manning, J., Matchew, C., McCarthy, M.,
              Meldrim, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J.,
              Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
              O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
              Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
              Roman, J., Schauer, S., Schupack, R., Seaman, S., Severy, P., Smith, C.,
              Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
              Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M.,
              Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X.,
              Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
              Direct Submission
              Submitted (21-MAR-2003) Whitehead Institute/MIT Center for Genome
              Research, 320 Charles Street, Cambridge, MA 02141, USA
              On Mar 21, 2003 this sequence version replaced gi:2877278.
              All repeats were identified using RepeatMasker:
              Smt, A.F.A. & Green, P. (1996-1997)
              http://ftp.genome.washington.edu/RM/RepeatMasker.html
              ----- Genome Center
              Center: Whitehead Institute/ MIT Center for Genome Research
              Center code: MIBR
              Web site: http://www-seq.wi.mit.edu
              Contact: sequence_submissions@genome.wi.mit.edu
              ----- Project Information
              Center project name: 127824
              Center clone name: 860_O_1
              -----
              Location/Qualifiers
              1..214946
              /organism="Homo sapiens"
              /mol_type="genomic DNA"
              /db_xref="taxon:9606"
              /chromosome="15"
              /map="15"
              /clone="RP11-86001"
              /clone_11b="RP11-11 Human Male BAC"
              /complement(1..362)
              /rpc_family="HALL1"
              /complement(823..1431)
              /rpc_family="LIMCS"
              913..922
              /note="<30 qual SNGL region"
              /complement(2941..3295)
              /rpc_family="MER31B"
              /complement(3307..3311)
              /note="<30 qual SNGL region"
              3438..3476
              /note="single clone coverage"
              /note="GA-rich"
              3546..3586
              /rpc_family="GA-rich"
              3601..4006
              /rpc_family="MT1A1"
              /complement(4008..4166)
              /rpc_family="MIR"
              /complement(4580..4711)

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
Roman, J., Schauer, S., Schupack, R., Seaman, S., Severy, P., Smith, C.,
Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M.,
Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (23-JAN-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 214946)
Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
Anderson, S., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T.,
Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y.,
Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K.,
Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S.,
Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hago, B.,
Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C.,
Macdonald, P., Major, J., Manning, J., Matchew, C., McCarthy, M.,
Meldrim, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J.,
Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Schauer, S., Schupack, R., Seaman, S., Severy, P., Smith, C.,
Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M.,
Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (21-MAR-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 21, 2003 this sequence version replaced gi:2877278.
All repeats were identified using RepeatMasker:
Smt, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: 127824
Center clone name: 860_O_1
-----
Location/Qualifiers
1..214946
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="15"
/map="15"
/clone="RP11-86001"
/clone_11b="RP11-11 Human Male BAC"
/complement(1..362)
/rpc_family="HALL1"
/complement(823..1431)
/rpc_family="LIMCS"
913..922
/note="<30 qual SNGL region"
/complement(2941..3295)
/rpc_family="MER31B"
/complement(3307..3311)
/note="<30 qual SNGL region"
3438..3476
/note="single clone coverage"
/note="GA-rich"
3546..3586
/rpc_family="GA-rich"
3601..4006
/rpc_family="MT1A1"
/complement(4008..4166)
/rpc_family="MIR"
/complement(4580..4711)

```



```

repeat_region      /rpt_family="L3"
4731..4751
/rpt_family="AT_rich"
repeat_region      complement(4758..4993)
/rpt_family="MIR"
repeat_region      complement(5336..5464)
/rpt_family="L3"
repeat_region      5935..6412
/rpt_family="MER39"
repeat_region      complement(6466..7069)
/rpt_family="MLT1B"
repeat_region      complement(8633..8713)
/rpt_family="MIR"
repeat_region      8768..9244
/rpt_family="L1MA2"
repeat_region      complement(9299..9498)
/rpt_family="MIR"
repeat_region      10257..10587
/rpt_family="MLT1A0"
repeat_region      12232..12619
/rpt_family="L1MA5"
repeat_region      complement(13183..13354)
/rpt_family="MIR"
repeat_region      14074..14500
/rpt_family="L1MB3"
repeat_region      15070..15201
/rpt_family="FLAM_C"
repeat_region      complement(16057..18621)
/rpt_family="L1MB3"
repeat_region      complement(18678..19377)
/rpt_family="L1M1"
repeat_region      complement(19375..20142)
/rpt_family="L1MB3"
repeat_region      22615..22919
/rpt_family="A1USg"
repeat_region      complement(25101..25206)
/rpt_family="HAL1"
repeat_region      complement(25691..25925)
/rpt_family="HAL1"
repeat_region      25996..26033
/rpt_family="CAGA)n"
repeat_region      complement(26059..26582)
/rpt_family="MLT2B3"
repeat_region      26645..27455
/rpt_family="MER21B"
repeat_region      complement(27582..27833)
/rpt_family="HAL1"
repeat_region      complement(28185..28759)
/rpt_family="L1TR68"
repeat_region      complement(29229..29603)
/rpt_family="THE1A"
repeat_region      complement(29771..30036)
/rpt_family="A1uJb"
repeat_region      30056..30083
/rpt_family="(TTTA)n"
repeat_region      complement(30090..30369)
/rpt_family="A1uJo"
repeat_region      31042..32585
/rpt_family="L1M4"
repeat_region      32586..32654
/rpt_family="(TCTA)n"
repeat_region      32655..33163
/rpt_family="L1M4"
repeat_region      33164..33473
/rpt_family="A1uY"
repeat_region      33474..33683
/rpt_family="L1M4"
repeat_region      33686..33751
/rpt_family="MLT1D"
repeat_region      36941..37015
/rpt_family="MIR"
repeat_region      complement(38001..38449)
/rpt_family="L1MD1"

```

```

repeat_region      38453..38855
/rpt_family="MLT1D"
repeat_region      complement(38856..39656)

Query Match          21.9%; Score 34.4; DB 9; Length 214946;
Best Local Similarity 54.8%; Pred. No. 3.7;
Matches 68; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY      5 TCATCTCAGATGCGATACATGCTGTTCACCCGTCCTGTGTCATGTCGCAAGC 64
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      35745 TCATCTCTATGGGACAGAAAAATCATTTGACAAATATGACATCTTCATGATGTAAC 35804
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY      65 CTGCGCTATTTGACCAAGAGATACCTACCTCCCAAGAACATCACTACTATGCAAC 124
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      35805 ATCCACATATCATGAAAGAAAGATGCTTTCTTAACCAATAAATGATACCTATGAAAA 35864
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY      125 TTCC 128
|||||
Db      35865 ATCC 35868

RESULT 14
CR847532/c 141810 bp DNA linear HTG 03-OCT-2004
LOCUS
DEFINITION
DANIO rerio clone DKRYP-86C9, *** SEQUENCING IN PROGRESS ***, 8
unordered pieces.
ACCESSION
CR847532
VERSION
CR847532.2 GI:53755907
KEYWORDS
HTG; HTGS PHASE1.
SOURCE
Danio rerio (zebrafish)
ORGANISM
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 141810)
McLay K.
Direct Submission
Submitted (02-OCT-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
fish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Oct 3, 2004 this sequence version replaced gi:53145936.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: zK086C9
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 139909 bases at least Q40
Consensus quality: 140415 bases at least Q30
Consensus quality: 140735 bases at least Q20
Insert size: 141110; sum-of-contigs
Insert size: 140600; 2.5% error; agarose-fp
Quality coverage: 8.52x in Q20 bases; sum-of-contigs Quality
coverage: 8.64x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 19172: contig of 19172 bp in length
* 19173 19272: gap of 100 bp
* 19273 46769: contig of 27497 bp in length
* 46770 46869: gap of 100 bp
* 46870 52415: contig of 5546 bp in length
* 52416 52515: gap of 100 bp

```

```
* 52516 64328: config of 11813 bp in length
* 64429 64428: gap of 100 bp
* 64429 96324: config of 31796 bp in length
* 96325 96324: gap of 100 bp
* 96325 121956: config of 25632 bp in length
* 121957 122056: gap of 100 bp
* 122057 131332: config of 9276 bp in length
* 131333 131432: gap of 100 bp
* 131433 141810: config of 10378 bp in length.
```

FEATURES

source

```
1. 141810
  /organism="Danio rerio"
  /mol_type="genomic DNA"
  /db_xref="taxon:7955"
  /clone="DKEYP-86C9"
  /clone_11b="DanioKeyplot"
  1. 19172
  /note="assembly: fragment:00300
  fragment:chain:1"
  19273. 46769
  /note="assembly: fragment:00469
  fragment:chain:1"
  46870. 52415
  /note="assembly: fragment:00017
  fragment:chain:1"
  52516. 64328
  /note="assembly: fragment:00167
  fragment:chain:1"
  64429. 96324
  /note="assembly: fragment:00976
  fragment:chain:2"
  96325. 121956
  /note="assembly: fragment:01302
  fragment:chain:2"
  122057. 131332
  /note="assembly: fragment:00056.0"
  131433. 141810
  /note="assembly: fragment:00703"
```

ORIGIN

```
Query Match 21.5%; Score 33.8; DB 2; Length 141810;
Best Local Similarity 52.5%; Pred. No. 5.8;
Matches 74; Conservative 0; Mismatches 67; Indels 0; Gaps 0;
```

```
QY 12 ACATGGCATACTACATGCTGTTCACCGTTCCTGTTGTCATGCAAGCTTGCGCT 71
   |||
Db 41846 ACAGTACCCCTACATCTTGTGCTCCTCAATGCAATTTGTAACAAGTACAGAGTTGGGT 41787
   |||
QY 72 ATTCTGAACCAAGAGATACCTACTCCCAAAATCCATCTTACTCATGCAACTTCCATG 131
   |||
Db 41786 AAGTGGGAATATGGGATCATCTGCTCTGAAAAATGAAATATATACACAAGAAACTT 41727
   |||
QY 132 CAACACGACATATGTTTCC 152
   |||
Db 41726 TCAAAAAAACTTATTAGTCC 41706
   |||
```

RESULT 15

```
EX640577 214178 bp DNA linear VRT 10-UTL-2004
LOCUS Zebrafish DNA sequence from clone DKEY-117014 in linkage group 14,
DEFINITION complete sequence.
ACCESSION BX640577
VERSION BX640577.17 GI:50080105
KEYWORDS
```

SOURCE

ORGANISM

Danio rerio (zebrafish)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Ostariophysi;

Cypriniformes; Cyprinidae; Danio.

REFERENCE 1 (bases 1 to 214178)

AUTHORS McGilre,S.

TITLE Direct Submission

JOURNAL

COMMENT

Submitted (10-JUN-2004) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfish-help@sanger.ac.uk Clone requesters: clonerequest@sanger.ac.uk On Jul 8, 2004 this sequence version replaced GI:49659408.

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: <http://www.sanger.ac.uk>

Contact: zfish-help@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information

on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep/Clone-derived

zebrafish pUC subclones occasionally display inconsistency over the

length of mononucleotide A/T runs and conserved TA repeats. Where

this is found the longest good quality representation will be

submitted.

Repeat names beginning 'Dr' were identified by the Recon repeat

discovery system (Zhifeng Bao and Sean Eddy, submitted), and those

beginning 'dr' were identified by Rick Waterman (Stephen Johnson

lab, Washu). For further information see

http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml DKEY-117014

is from a Zebrafish BAC library

VECTOR: pindigOBAC-5

FEATURES

source

```
1. 214178
  /organism="Danio rerio"
  /mol_type="genomic DNA"
  /db_xref="taxon:7955"
  /clone="DKEY-117014"
  /clone_11b="DanioKey"
```

ORIGIN

```
Query Match 21.5%; Score 33.8; DB 5; Length 214178;
Best Local Similarity 52.5%; Pred. No. 5.9;
Matches 74; Conservative 0; Mismatches 67; Indels 0; Gaps 0;
```

```
QY 12 ACATGGCATACTACATGCTGTTCACCGTTCCTGTTGTCATGCAAGCTTGCGCT 71
   |||
Db 58301 ACAGTACCCCTACATCTTGTGCTCCTCAATGCAATTTGTAACAAGTACAGAGTTGGGT 58360
   |||
QY 72 ATTCTGAACCAAGAGATACCTACTCCCAAAATCCATCTTACTCATGCAACTTCCATG 131
   |||
Db 58361 AAGTGGGAATATGGGATCATCTGCTCTGAAAAATGAAATATATACACAAGAAACTT 58420
   |||
QY 132 CAACACGACATATGTTTCC 152
   |||
Db 58421 TCAAAAAAACTTATTAGTCC 58441
   |||
```

Search completed: September 16, 2005, 03:01:36
Job time : 528.823 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 08:10:17 ; Search time 76.0101 Seconds
(without alignments)
12227.321 Million cell updates/sec

Title: US-10-713-381-2_COPY_1155_1311

Perfect score: 157
Sequence: 1 cgtgcacatccacatgcat.....cgcacatcgtcttcctgaac 157

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneeqn19808:*
2: geneeqn19908:*
3: geneeqn20008:*
4: geneeqn20018:*
5: geneeqn20028:*
6: geneeqn20038:*
7: geneeqn20048:*
8: geneeqn20058:*
9: geneeqn20068:*
10: geneeqn20078:*
11: geneeqn20088:*
12: geneeqn20098:*
13: geneeqn20108:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	157	100.0	255	5	AAH76340	Aah76340 Z. mays M
2	157	100.0	1394	2	AAH07408	Aax07408 Zea mays
3	157	100.0	1394	2	AAH07409	Aax07409 Zea mays
4	157	100.0	1394	5	AAH76332	Aah76332 Z. mays M
5	157	100.0	1394	5	AAH76333	Aah76333 Z. mays M
6	146	93.0	158	5	AAH76334	Aah76334 Z. mays M
7	50	31.8	50	5	AAH76336	Aah76336 Z. mays M
8	40	25.5	40	5	AAH76337	Aah76337 Z. mays M
9	32.4	20.6	3257	12	ADK52331	Adk52331 Mouse atcc
c 10	31.6	20.1	2352	8	ACA48402	Ac48402 Prokaryot
c 11	31.4	20.0	15734	6	ABN80329	Abn80329 Human che
12	30.6	19.5	64415	3	AAF22279	Aaf22279 BAC conta
13	30.6	19.5	79467	9	ADA02717	Ada02717 Mouse Nfa
14	30.6	19.5	79467	10	ADB92455	Adb92455 Mouse Nfa
15	30.6	19.5	79467	10	ADK95965	Adk95965 Mouse Nfa
16	30.4	19.4	52001	12	ADP67076	Adp67076 Human nfa
c 17	30.2	19.2	669	3	AAAS1436	Aaas1436 A. thalia
18	30.2	19.2	1266	9	ADB08983	Adb08983 Allotococ
19	30.2	19.2	110000	9	ADB12064_07	Continuation (8 of
20	30	19.1	30	5	AAH76335	Aah76335 Z. mays M

c	21	30	19.1	378	10	ADD33403	Add33403 Mouse mit
c	22	30	19.1	396	10	ADD33404	Add33404 Mouse mit
c	23	29.8	19.0	373	4	AAH13328	Aah13328 Human cdn
c	24	29.6	18.9	1185	10	ADF00796	Adf00796 Bacterial
c	25	29.4	18.7	110000	3	AAF22303_2	Continuation (3 of
c	26	29.2	18.6	1475	9	AAI62902	Aai62902 Vernonia
c	27	29.2	18.6	110000	2	AAH76333_06	Continuation (7 of
c	28	28.8	18.3	1500	8	ADA68284	Ada68284 Arabidope
c	29	28.6	18.2	756	13	ADS57339	Ads57339 Bacterial
c	30	28.6	18.2	778	4	AAH32892	Aah32892 Human gen
c	31	28.6	18.2	1043	3	AAH32893	Aah32893 Human gen
c	32	28.6	18.2	1043	3	ADD71973	Add71973 Human urt
c	33	28.6	18.2	1043	3	ABN67847	Abn67847 Streptoco
c	34	28.6	18.2	51615	13	ACN45162	Acn45162 Human gen
c	35	28.6	18.2	110000	13	ACN45162_5	Continuation (6 of
c	36	28.6	18.2	164702	8	ACF62730	Acf62730 Cancer ba
c	37	28.6	18.2	164702	8	ADB20845	Adb20845 KRP1 base
c	38	28.6	18.2	164702	10	ADB87934	Adb87934 Human UGT
c	39	28.6	18.2	164702	10	ADB96917	Adb96917 Human MDR
c	40	28.6	18.2	164702	10	ADB92108	Adb92108 Human MDR
c	41	28.4	18.1	1478	3	AAH15843	Aah15843 Human pio
c	42	28.4	18.1	1715	8	ABA00762	AbA00762 PCR tela
c	43	28.4	18.1	1791	12	ADQ86652	Adq86652 Human tum
c	44	28.4	18.1	1791	12	ADQ85536	Adq85536 Human tum
c	45	28.4	18.1	1791	12	ADQ84654	Adq84654 Human tum

ALIGNMENTS

RESULT 1
AAH76340
ID AAH76340 standard; DNA; 255 BP.
XX
AC AAH76340;
XX
DT 29-OCT-2001 (first entry)
XX
DE Z. mays Ms45 promoter fragment.
XX
KW Ms45; male tissue; regulatory region; transcription; male fertility;
KM hybrid seed; promoter; de.
XX
OS Zea mays.
XX
PN WO200160997-A2.
XX
PD 23-AUG-2001.
XX
PP 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
DR WPI; 2001-514772/56.
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
XX essential for initiating transcription of the Ms45 gene useful for
XX mediating fertility in a male plant.
XX
PS Example 5; Fig 8; 50bp; English.
XX
XX The invention provides a male tissue-preferred regulatory region (I)
XX comprising nucleotide sequences essential for initiating transcription of
XX the Ms45 gene. A method of mediating male fertility in a plant is
XX provided that involves introducing an expression vector comprising a
XX promoter operably linked to (I) into a plant where the exogenous gene
XX impacts male fertility of the plant and (II) controls expression of the
XX exogenous gene. A method of producing hybrid seeds is also provided. The
XX present sequence represents a Z. mays Ms45 promoter fragment

```
XX Sequence 255 BP; 59 A; 86 C; 39 G; 71 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 157; DB 5; Length 255;
Best Local Similarity 100.0%; Pred. No. 1.3e-42;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGTGCATCTACATGGAATGATGCTGTTGCAACCGTTGCTGTTGCTGATGCTCC 60
DB 15 CGTGCATCTACATGGAATGATGCTGTTGCAACCGTTGCTGTTGCTGATGCTCC 74
QY 61 AAGCCTTGCTATTCTGAAACCAAGAGATTAATCTCCCAACATCTTACTGATG 120
DB 75 AAGCCTTGCTATTCTGAAACCAAGAGATTAATCTCCCAACATCTTACTGATG 134
QY 121 CAACCTTCATGCAACACGACATATGTTTCTGTAAC 157
DB 135 CAACCTTCATGCAACACGACATATGTTTCTGTAAC 171

RESULT 2
AA07408 standard; DNA; 1394 BP.
XX
AC AAX07408;
XX
DT 08-JUN-1999 (first entry)
XX
DE Zea mays M645 male tissue-preferred regulatory region.
XX
KW M645; male; tissue-preferred; regulatory region; plant cells;
KW plant tissue; differentiated; maize; hybrid seed; fertility; ss.
XX
OS Zea mays.
XX
XX WO9859061-A1.
XX
XX 30-DEC-1998.
XX
XX 19-JUN-1998; 98WO-US012895.
XX
XX 23-JUN-1997; 97US-00880499.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX PI Albertsen MC, Fox TW, Garnaat CW, Huffman GA, Kendall TL;
XX
XX WPI; 1999-105628/09.
XX
XX New nucleic acid encoding a M645 male tissue-preferred regulatory region
XX - useful in mediating plant fertility, especially hybrid seed production.
XX
XX Claim 2; Page 22-23; 39pp; English.
XX
XX The sequence is that encoding an M645 male tissue-preferred regulatory
XX region. It may be used in the construction of a vector for a method of
XX producing exogenous genes in a male tissue-preferred manner, which is
XX useful in restoring or conferring fertility, such as in hybrid seed
XX production. In conferring fertility, a monocot/dicot plant is transformed
XX with the exogenous nucleotide sequence (a male sterility gene, preferably
XX M645), which encodes a product selected from auxins, rolls and diptheria
XX toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
XX and infertile plants
XX
XX Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 157; DB 2; Length 1394;
Best Local Similarity 100.0%; Pred. No. 2.1e-42;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGTGCATCTACATGGAATGATGCTGTTGCAACCGTTGCTGTTGCTGATGCTCC 60
DB 1155 CGTGCATCTACATGGAATGATGCTGTTGCAACCGTTGCTGTTGCTGATGCTCC 1214
```

```
QY 61 AAGCCTTGCTATTCTGAAACCAAGAGATTAATCTCCCAACATCTTACTGATG 120
DB 1215 AAGCCTTGCTATTCTGAAACCAAGAGATTAATCTCCCAACATCTTACTGATG 1274
QY 121 CAACCTTCATGCAACACGACATATGTTTCTGTAAC 157
DB 1275 CAACCTTCATGCAACACGACATATGTTTCTGTAAC 1311

RESULT 3
AA07409 standard; DNA; 1394 BP.
XX
AC AAX07409;
XX
DT 08-JUN-1999 (first entry)
XX
DE Zea mays M645 male tissue-preferred regulatory region.
XX
KW M645; male; tissue-preferred; regulatory region; plant cells;
KW plant tissue; differentiated; hybrid seed; fertility; ss.
XX
OS Zea mays.
XX
XX WO9859061-A1.
XX
XX 30-DEC-1998.
XX
XX 19-JUN-1998; 98WO-US012895.
XX
XX 23-JUN-1997; 97US-00880499.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX PI Albertsen MC, Fox TW, Garnaat CW, Huffman GA, Kendall TL;
XX
XX WPI; 1999-105628/09.
XX
XX New nucleic acid encoding a M645 male tissue-preferred regulatory region
XX - useful in mediating plant fertility, especially hybrid seed production.
XX
XX Claim 3; Page 23-24; 39pp; English.
XX
XX The sequence is that encoding an M645 male tissue-preferred regulatory
XX region. It may be used in the construction of a vector for a method of
XX producing exogenous genes in a male tissue-preferred manner, which is
XX useful in restoring or conferring fertility, such as in hybrid seed
XX production. In conferring fertility, a monocot/dicot plant is transformed
XX with the exogenous nucleotide sequence (a male sterility gene, preferably
XX M645), which encodes a product selected from auxins, rolls and diptheria
XX toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
XX and infertile plants
XX
XX Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 157; DB 2; Length 1394;
Best Local Similarity 100.0%; Pred. No. 2.1e-42;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGTGCATCTACATGGAATGATGCTGTTGCAACCGTTGCTGTTGCTGATGCTCC 60
DB 1155 CGTGCATCTACATGGAATGATGCTGTTGCAACCGTTGCTGTTGCTGATGCTCC 1214
QY 61 AAGCCTTGCTATTCTGAAACCAAGAGATTAATCTCCCAACATCTTACTGATG 120
DB 1215 AAGCCTTGCTATTCTGAAACCAAGAGATTAATCTCCCAACATCTTACTGATG 1274
QY 121 CAACCTTCATGCAACACGACATATGTTTCTGTAAC 157
DB 1275 CAACCTTCATGCAACACGACATATGTTTCTGTAAC 1311
```

RESULT 4
AAH76332 standard; DNA; 1394 BP.
XX
AC AAH76332;
XX
DT 29-OCT-2001 (first entry)
XX
DE Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.
XX
KM Ms45; male tissue; regulatory region; transcription; male fertility;
XX
KM hybrid seed; ds.
XX
OS Zea mays.
XX
PN WO200160997-A2.
XX
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albrechtsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
DR WPI; 2001-514772/56.
XX
XX A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the Ms45 gene useful for
PT mediating fertility in a male plant.
XX
XX Claim 4; Page 46; 50pp; English.
XX
PS The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the Ms45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a nucleic acid sequence encoding an Ms45 male
CC tissue preferred regulatory region from Z. mays
SQ Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;
Query Match 100.0%; Score 157; DB 5; Length 1394;
Best Local Similarity 100.0%; Pred. No. 2.1e-42;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGTGTCTCATCTCAGATGAGCATCTAGTCTGTTCAACGGTTCGTTGTTCCATGCTCC 60
DB 1155 CGTGTCTCATCTCAGATGAGCATCTAGTCTGTTCAACGGTTCGTTGTTCCATGCTCC 1214
QY 61 AAGCCTTGCTATTCTGAACCAAGAGATACCTACTCCCAACCAATCATCTTACTCATG 120
DB 1215 AAGCCTTGCTATTCTGAACCAAGAGATACCTACTCCCAACCAATCATCTTACTCATG 1274
QY 121 CAACTTCATGCAAAACGACGACATATGTTTCTGTAAC 157
DB 1275 CAACTTCATGCAAAACGACGACATATGTTTCTGTAAC 1311
RESULT 5
AAH76333 standard; DNA; 1394 BP.
XX
AC AAH76333;
XX
DT 29-OCT-2001 (first entry)
XX
DE Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.

XX
KM Ms45; male tissue; regulatory region; transcription; male fertility;
XX
KM hybrid seed; ds.
XX
OS Zea mays.
XX
PN WO200160997-A2.
XX
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albrechtsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
DR WPI; 2001-514772/56.
XX
XX A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the Ms45 gene useful for
PT mediating fertility in a male plant.
XX
XX Claim 4; Page 47; 50pp; English.
XX
PS The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the Ms45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a nucleic acid sequence encoding an Ms45 male
CC tissue preferred regulatory region from Z. mays
SQ Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;
Query Match 100.0%; Score 157; DB 5; Length 1394;
Best Local Similarity 100.0%; Pred. No. 2.1e-42;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGTGTCTCATCTCAGATGAGCATCTAGTCTGTTCAACGGTTCGTTGTTCCATGCTCC 60
DB 1155 CGTGTCTCATCTCAGATGAGCATCTAGTCTGTTCAACGGTTCGTTGTTCCATGCTCC 1214
QY 61 AAGCCTTGCTATTCTGAACCAAGAGATACCTACTCCCAACCAATCATCTTACTCATG 120
DB 1215 AAGCCTTGCTATTCTGAACCAAGAGATACCTACTCCCAACCAATCATCTTACTCATG 1274
QY 121 CAACTTCATGCAAAACGACGACATATGTTTCTGTAAC 157
DB 1275 CAACTTCATGCAAAACGACGACATATGTTTCTGTAAC 1311
RESULT 6
AAH76334 standard; DNA; 158 BP.
XX
AC AAH76334;
XX
DT 29-OCT-2001 (first entry)
XX
DE Z. mays Ms45 male tissue-preferred regulatory region fragment.
XX
KM Ms45; male tissue; regulatory region; transcription; male fertility;
XX
KM hybrid seed; ds.
XX
OS Zea mays.
XX
PN WO200160997-A2.
XX
PD 23-AUG-2001.

XX 13-FEB-2001; 2001WO-US004527.
PF
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
XX WPI; 2001-514772/56.
DR
XX
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the MS45 gene useful for
PT mediating fertility in a male plant.
XX
PS Claim 5; Page 47; 50pp; English.
XX
XX The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the MS45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a DNA fragment -38 to -195 bases upstream of
CC the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
CC nucleotide sequence
XX
SO Sequence 158 BP; 41 A; 50 C; 21 G; 46 T; 0 U; 0 Other;
XX
Query Match 93.0%; Score 146; DB 5; Length 158;
Best Local Similarity 99.4%; Pred. No. 5.4e-39;
Matches 157; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
XX
QY 1 CGGTGATCTCAATGCGATCTACATGCTTGTTCACCGTTGTC-TTGTTCATCGTC 59
DB 1 CGGTGATCTCAATGCGATCTACATGCTTGTTCACCGTTGTC-TTGTTCATCGTC 60
XX
QY 60 CAAGCCTTGCTTCTTGAACCAAGAGATACCTACTCCCAACATCCATTACTCAT 119
DB 61 CAAGCCTTGCTTCTTGAACCAAGAGATACCTACTCCCAACATCCATTACTCAT 120
XX
QY 120 GCAACTTCCATGCAACACGACATATGTTTCTGTAAC 157
DB 121 GCAACTTCCATGCAACACGACATATGTTTCTGTAAC 158
XX
RESULT 7
AAH76336
ID AAH76336 standard; DNA; 50 BP.
XX
AC AAH76336;
XX
DT 29-OCT-2001 (first entry)
XX
DE 2. mays Ms45 male tissue-preferred regulatory region fragment.
XX
XX Ms45; male tissue; regulatory region; transcription; male fertility;
KM hybrid seed; ds.
XX
OS Zea mays.
XX
XX WO200160997-A2.
XX
XX PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;

XX
DR WPI; 2001-514772/56.
XX
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the MS45 gene useful for
PT mediating fertility in a male plant.
XX
PS Claim 5; Page 47; 50pp; English.
XX
XX The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the MS45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a DNA fragment -72 to -111 bases upstream of
CC the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
CC nucleotide sequence
XX
SQ Sequence 50 BP; 18 A; 17 C; 5 G; 10 T; 0 U; 0 Other;
XX
Query Match 31.8%; Score 50; DB 5; Length 50;
Best Local Similarity 100.0%; Pred. No. 7.2e-07;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 75 CTGAACCAAGAGATACCTACTCCCAACATCCATTACTGACGAAC 124
DB 1 CTGAACCAAGAGATACCTACTCCCAACATCCATTACTGACGAAC 50
XX
RESULT 8
AAH76337
ID AAH76337 standard; DNA; 40 BP.
XX
AC AAH76337;
XX
DT 29-OCT-2001 (first entry)
XX
DE 2. mays Ms45 male tissue-preferred regulatory region fragment.
XX
KM Ms45; male tissue; regulatory region; transcription; male fertility;
KW hybrid seed; ds.
XX
OS Zea mays.
XX
XX WO200160997-A2.
XX
XX PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
DR WPI; 2001-514772/56.
XX
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the MS45 gene useful for
PT mediating fertility in a male plant.
XX
PS Claim 14; Page 32; 50pp; English.
XX
XX The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the MS45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The

CC Present sequence represents a DNA fragment upstream of the TATA box of a
CC Z. mays M645 male-tissue preferred regulatory region nucleotide sequence
XX
SQ Sequence 40 BP; 14 A; 14 C; 3 G; 9 T; 0 U; 0 Other;
Query Match 25.5%; Score 40; DB 5; Length 40;
Best Local Similarity 100.0%; Pred. No. 0.0015;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 85 AGGATACCTACTCCCAACAATCCATCTTACTCATGCAAC 124
DB 1 AGGATACCTACTCCCAACAATCCATCTTACTCATGCAAC 40
RESULT 9
ADK52131
ID ADK52131 standard; cDNA; 3267 BP.
XX
AC ADK52131;
XX
DT 20-MAY-2004 (first entry)
XX
DE Mouse atopic-dermatitis/psoriasis-associated EST #4.
XX
KM Mouse; BS; EST; atopic dermatitis; psoriasis; dermatological;
KM anti-inflammatory; antipsoriatic; rash; expressed sequence tag.
XX
OS Mus musculus.
XX
PN MO2004016785-A1.
XX
PD 26-FEB-2004.
XX
PF 06-AUG-2003; 2003MO-JP009999.
XX
PR 06-AUG-2002; 2002JP-00229319.
PR 14-MAY-2003; 2003JP-00136544.
XX
PA (GENO-) GENOX RES INC.
PA (UYJU-) UNIV JUNTENDO.
XX
PI Itoh M, Ogawa K, Shinagawa A, Sudo H, Ogawa H, Ra C;
PI Mitsutshiri K;
DR MPI; 2004-214514/20.
XX
PT Detecting atopic dermatitis or psoriasis comprises assaying levels of
PT expression of an indicator gene at a rash site and non-rash site of a
PT person with atopic dermatitis or psoriasis.
XX
PS Claim 20; SEQ ID NO 164; 484bp; Japanese.
XX
CC The invention relates to detecting atopic dermatitis or psoriasis
CC comprising assaying the levels of expression of an indicator gene at a
CC rash site and non-rash site of a person with atopic dermatitis or
CC psoriasis, comprising these levels with those of a healthy person, and
CC determining that if the levels of indicators are higher or lower, then
CC this indicates the disease. Also included are a reagent for detecting
CC atopic dermatitis or psoriasis, a kit for screening for treatments, a
CC transgenic non human vertebrate animal models for the diseases, an agent
CC for inducing the diseases in mice and a DNA chip for assaying for the
CC indicator genes. The method is used for treatment, detection and animal
CC models for research of atopic dermatitis and psoriasis. The present
CC sequence is a Mouse atopic-dermatitis/psoriasis-associated EST
CC (expressed sequence tag).
SQ Sequence 3267 BP; 1029 A; 658 C; 590 G; 990 T; 0 U; 0 Other;
Query Match 20.6%; Score 32.4; DB 12; Length 3267;
Best Local Similarity 54.1%; Pred. No. 2.2;
Matches 66; Conservative 0; Mismatches 56; Indels 0; Gaps 0;
QY 27 TGCTGTTCACACCGTTCGTCTGTTCATCGTCACAGCTTGCTATTCGAACCAAGAG 86

DB 2774 TTCTTTTTCACCTCCCTCTGATTTCTTTGTAACCTCACCTTCATTTTGCCCTTTAAGTT 2833
QY 87 GATACCTACTCCCAACAATCCATCTTACTCATGCAACTTCCACGCAACATAT 146
DB 2834 GTTCTGATTTCCCAAGAAAGCCATTTTGACATGACGTCAAAAGATATAACCTGAAAT 2893
QY 147 GT 148
DB 2894 GT 2895
RESULT 10
ACA48402/C
ID ACA48402 standard; DNA; 2352 BP.
XX
AC ACA48402;
XX
DT 19-JUN-2003 (first entry)
XX
DE Prokaryotic essential gene #30059.
XX
KM Antisense; ds; prokaryotic essential gene; cell proliferation;
KM drug design; gene.
XX
OS Streptococcus mutans.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002MO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362899P.
XX
PA (ELITR-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zykkind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR MPI; 2003-029926/02.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 14; SEQ ID NO 36272; 1766bp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antipositive; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of

CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 2352 BP; 762 A; 401 C; 500 G; 689 T; 0 U; 0 Other;

Query Match 20.1%; Score 31.6; DB 8; Length 2352;

Best Local Similarity 58.5%; Pred. No. 3.8;

Matches 55; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 62 AGCCTTGCTATTCTGAACCAAGAGATACCTACTCCCAACAATCCTTACTGATGC 121

Db 234 ATCTGACTGTGCCCAAGCAAGAGACATCTCTGAAAATGTGCTACATCTGC 175

QY 122 AACTTCATGCAACACGACATATGTTCTCTGA 155

Db 174 AATTTCAGCAAAAGAACGTTCAATTTTCTCTGA 141

RESULT 11

ABN80329/C

ID ABN80329 standard; DNA; 15734 BP.

XX AC ABN80329;

XX DT 15-JUL-2002 (first entry)

XX DE Human chemically modified disease associated gene SEQ ID NO 346.

XX DE Human; development; homeobox gene; HOX; diabetes; cancer; apoptosis;

XX KM heart disease; epilepsy; histone deacetylation; muscular dystrophy;

XX KM dwarfism; single nucleotide polymorphism; SNP; cytosine methylation;

XX KM antidiabetic; cytostatic; anticonvulsant; ds.

XX OS Homo sapiens.

XX OS Synthetic.

XX PN WO200200927-A2.

XX PD 03-JAN-2002.

XX PF 02-JUL-2001; 2001WO-EP007536.

XX PR 30-JUN-2000; 2000DE-01032529.

XX PR 01-SEP-2000; 2000DE-01043826.

XX PA (EPig-) EPIGENOMICS AG.

XX PI Olek A, Piepenbrock C, Berlin K;

XX DR WPI; 2002-130908/17.

XX PT Novel nucleic acid useful for diagnosis and therapy of diseases

XX PT associated with development genes such as diabetes, comprises a sequence

XX PT of a segment of chemically pretreated DNA of genes associated with

XX PT development.

XX PS Claim 1; SEQ ID NO 346; 27pp; English.

XX CC The invention relates to a nucleic acid (I) comprising a sequence at

XX CC least 18 bases in length of a segment of chemically pretreated DNA (II)

XX CC of genes associated with development selected from 87 genes listed in the

XX CC specification such as ACCPN, ADFN, or AFDI and comprising one of 350

XX CC sequences (ABN79984-ABN80333) or their complements. The invention is

XX CC useful for the diagnosis or therapy of diseases associated with

CC development genes, in particular disease related to homeobox containing
CC genes (HOX), like diabetes, cancer, apoptosis related diseases, syndromes
CC associated with congenital heart disease, epilepsy, diseases related to
CC histone deacetylation, Currarino syndrome, diseases related with the
CC development of the brain and limb girdle muscular dystrophy and dwarfism.
CC Oligomers specific to each of the genes are useful for detecting the
CC methylation state of all CpG dinucleotides within the 350 sequences or
CC (II) and their complementary sequences, as primer oligonucleotides for
CC the amplification of the 350 sequences, (II) and/or their complements and
CC as oligomer probes for detecting the cytosine methylation state and/or
CC single nucleotide polymorphisms (SNPs). Note: The sequence data for this
CC patent did not form part of the printed specification but is based on
CC sequence information supplied to Derwent by the European Patent Office

SQ Sequence 15734 BP; 3630 A; 744 C; 4234 G; 7126 T; 0 U; 0 Other;

Query Match 20.0%; Score 31.4; DB 6; Length 15734;

Best Local Similarity 59.6%; Pred. No. 8;

Matches 53; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 52 CGATGTCACAGCCTTGCTATTCTGAACCAAGAGATACCTACTCCCAACAATCCATC 111

Db 8802 CCTCTCTTCTCCCTTATCGAATCTTAACAATCAAAACCTTAATAATTAACACCAAC 8743

QY 112 TTACTCATGCAACTTCATGCAACACGCC 140

Db 8742 TACCTAACCCATCTTCCAAATTAACACC 8714

RESULT 12

AAF22279

ID AAF22279 standard; DNA; 64415 BP.

XX AC AAF22279;

XX DT 20-MAR-2001 (first entry)

XX DE BAC containing repeats from centromeres 1-4 #2.

XX KM Centromere; mitochondrion; vector; ds.

XX OS Arabidopsis thaliana.

XX OS WO200055325-A2.

XX PN 21-SEP-2000.

XX PD 17-MAR-2000; 2000WO-US007392.

XX PF 18-MAR-1999; 99US-0125219P.

XX PR 01-APR-1999; 99US-0127409P.

XX PR 18-MAY-1999; 99US-0134770P.

XX PR 13-SEP-1999; 99US-0153584P.

XX PR 17-SEP-1999; 99US-0154603P.

XX PR 16-DEC-1999; 99US-0172493P.

XX PA (UYCH-) UNIV CHICAGO.

XX PI Preuss D, Copenhaver G, Keith K;

XX DR WPI; 2000-587529/55.

XX PT Recombinant DNA construct comprising a plant centromere, useful for

XX PT producing stably inherited mitochondria which can serve as vectors for the

XX PT construction of transgenic plant and animal cells.

XX PS Claim 102; Page 321-335; 1449pp; English.

XX CC The present invention relates to a recombinant DNA construct of a plant

XX CC (Arabidopsis thaliana) centromere. The constructs are useful for

XX CC producing stably inherited mitochondria which can serve as vectors for the

XX CC construction of transgenic plant and animal cells expressing selected

XX CC proteins such as hormones, enzymes, interleukins, clotting factors,

CC cytokines, antibodies, and growth factors
XX
SQ Sequence 64415 BP, 18698 A, 13554 C, 13083 G, 18980 T, 0 U, 100 Other;
Query Match 19.5%; Score 30.6; DB 3; Length 64415;
Best Local Similarity 58.1%; Pred. No. 23;
Matches 54; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
QY 65 CTGGCCATCTGTGACCAAGAGTACTCTCCCAACCAATCCATTCTCTATGCAAC 124
DB 7117 CTGGCTCTTTAGATATATAGTATATCTCTCTAAACCTAAACCTAAACCTAACCC 7176
QY 125 TTCCATGCAACACGACATATATGTTCTCTGAC 157
DB 7177 TTAATCCCAACCTTAATCTAATCTCTTAAC 7209
RESULT 13
ADA02717
ID ADA02717 standard; DNA: 79467 BP.
XX
AC ADA02717;
XX
DT 06-NOV-2003 (first entry)
XX
DE Mouse Nfatc1 carcinoma associated gene, SEQ ID NO:1235.
XX
KM Mouse; murine; carcinoma associated; oncogene; carcinoma; cancer; breast;
KM prostate; lymphoma; leukemia; cytostatic; gene therapy; drug screening;
KM gene; ds.
XX
OS Mus sp.
XX
PN WO2003057146-A2.
XX
PD 17-JUL-2003.
XX
PF 26-DEC-2002; 2002WO-US041414.
XX
PR 26-DEC-2001; 2001US-00035832.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW;
XX
DR MPI; 2003-587068/55.
XX
PT New recombinant nucleic acid encoding carcinoma associated protein,
PT useful for preparing compositions for treating carcinomas.
XX
PS Claim 1; SEQ ID NO 1235; 245pp; English.
XX
CC The invention relates to recombinant carcinoma associated (CA) nucleic
CC acid sequences from mouse and human (ADA01482-AD03094), and to
CC recombinant carcinoma associated proteins (CAP) encoded by them. The
CC invention also encompasses expression vectors and host cells comprising a
CC CA nucleic acid, a polypeptide (especially an antibody) that specifically
CC binds to the protein, and a biochip comprising CA nucleic acid or
CC fragments thereof. The sequences of the invention were identified using
CC oncogenic retroviruses, which insert into the genome of the host organism
CC at random. Many of these do not carry transduced host oncogenes or
CC pathogenic trans-acting viral genes, meaning that cancer incidence is a
CC direct consequence of the effects of proviral integration into host
CC protooncogenes. The CA nucleic acid sequences can be used to diagnose
CC leukemia (especially breast cancer, prostate cancer, lymphoma or
CC leukemia) or a propensity to carcinoma by determination of the sequence
CC of a CA gene, or by determination of CA gene expression in particular
CC tissues. CA nucleic acids, proteins and antibodies are also useful as
CC therapeutic agents and in screening and evaluating drug candidates. The
CC present sequence represents a specifically claimed murine CA nucleic acid
CC sequence of the invention. Note: The complete sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 79467 BP, 18256 A, 19684 C, 20062 G, 20204 T, 0 U, 1261 Other;
Query Match 19.5%; Score 30.6; DB 9; Length 79467;
Best Local Similarity 56.4%; Pred. No. 25;
Matches 57; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
QY 51 TCCATGTCGACGCTTGCTTATTTGAAACCAAGAGATACCTATCCCAACATCCAT 110
DB 53779 TCCACTAGCTGATCTTGCTCTCTGGAACCAAGCTGAGCTTATGACACACACACA 53838
QY 111 CTTACTGATGCAACTTCGATCGAACAACGCAATATGTTTC 151
DB 53839 CACACACACACACACACACACACACACACACAGCTTGCTC 53879
RESULT 14
ADB72455
ID ADB72455 standard; DNA: 79467 BP.
XX
AC ADB72455;
XX
DT 04-DEC-2003 (first entry)
XX
DE Mouse Nfatc1 gene.
XX
KM mouse; ds; cytostatic; gene therapy; vaccine; carcinoma; lymphomas;
KM cancer; neoplasm; adenocarcinoma; sarcoma; gene.
XX
OS Mus sp.
XX
PN WO2003008583-A2.
XX
PD 30-JAN-2003.
XX
PF 26-DEC-2001; 2001WO-US051291.
XX
PR 02-MAR-2001; 2001US-00798586.
PR 23-OCT-2001; 2001US-00004113.
PR 08-NOV-2001; 2001US-00052482.
PR 30-NOV-2001; 2001US-00997722.
PR 20-DEC-2001; 2001US-00034650.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW, Engelhard EK;
XX
DR MPI; 2003-239337/23.
XX
PT New recombinant nucleic acid, useful for treating carcinomas, lymphomas,
PT cancers, neoplasm, adenocarcinoma, or sarcomas.
XX
PS Claim 1; SEQ ID NO 283; 2304pp; English.
XX
CC The invention relates to a novel recombinant nucleic acid comprising a
CC nucleotide sequence selected from any of the 660 sequences fully defined
CC in the specification. A polynucleotide of the invention has cytostatic
CC activity, and may have a use in gene therapy, or in a vaccine. The
CC recombinant nucleic acids and polypeptides are useful for treating
CC carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and
CC sarcomas. The present sequence represents a mouse gene of the invention.
XX
SQ Sequence 79467 BP, 18256 A, 19684 C, 20062 G, 20204 T, 0 U, 1261 Other;
XX
Query Match 19.5%; Score 30.6; DB 10; Length 79467;
Best Local Similarity 56.4%; Pred. No. 25;
Matches 57; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
QY 51 TCCATGTCGACGCTTGCTTATTTGAAACCAAGAGATACCTATCCCAACATCCAT 110
DB 53779 TCCACTAGCTGATCTTGCTCTCTGGAACCAAGCTGAGCTTATGACACACACACA 53838

QY 111 CTTACTCATGCACTTCCATGCAACAGCAGCATATGTTTC 151
 Db 53839 CACACACACACACACACACCAACACACAGCTTGTGGCTC 53879

RESULT 15

AD859565
 ID ADE95965 standard; DNA; 79467 BP.

XX ADE95965;

AC 12-FEB-2004 (first entry)

XX Mouse Nfatc1 gene genomic DNA sequence.

XX cancer diagnosis; cancer treatment; carcinoma; cytostatic; gene therapy;

KM lymphoma; breast cancer; prostate cancer; leukaemia; ds; mouse; murine;

XX Nfatc1.

OS Mus sp.

XX WO2003039484-A2.

XX 15-MAY-2003.

XX 08-NOV-2002; 2002WO-US036071.

XX 08-NOV-2001; 2001US-00052482.

XX (SAGR-) SAGRES DISCOVERY.

XX Morris DW, Engelhard EK;

XX WPI; 2003-441462/41.

XX New carcinoma associated nucleic acids and proteins, useful for screening

XX drug candidates, or for diagnosing and treating carcinomas, e.g.

XX lymphoma, breast cancer, prostate cancer or leukemia.

XX Claim 1; SEQ ID NO 223; 793pp; English.

XX This invention relates to novel recombinant nucleic acids for use in

XX diagnosis and treatment of cancer, especially carcinomas, as well as the

XX use of compositions in screening methods. The compositions of the

XX invention may have cytostatic activity whilst the disclosed sequences may

XX be useful for gene therapy. The carcinoma associated nucleic acids and

XX proteins are useful for diagnosing and treating carcinomas, for example

XX lymphoma, breast cancer, prostate cancer or leukemia, or for screening

XX drug candidates or bioactive agents capable of binding to, or modulating

XX the activity of, a carcinoma associated protein. The present sequence is

XX the genomic DNA sequence of the mouse Nfatc1 gene which is a carcinoma

XX associated gene of the invention.

XX SQ Sequence 79467 BP; 18256 A; 19684 C; 20062 G; 20204 T; 0 U; 1261 Other;

XX Query Match 19.5%; Score 30.6; DB 10; Length 79467;

XX Best Local Similarity 56.4%; Pred. No. 25;

XX Matches 57; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 51 TCCATGTCGAAGCCTTGCTATTTGAAACAAAGAGATACCTACTCCAAACATCCAT 110

Db 53779 TCCACTACTGATCCTTGCTCTCTGGAACAAAGCTGATGTCACACACACACA 53838

QY 111 CTTACTCATGCACTTCCATGCAACAGCAGCATATGTTTC 151

Db 53839 CACACACACACACACACACCAACACACAGCTTGTGGCTC 53879

Search completed: September 15, 2005, 21:33:21
 Job time : 77.0101 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 07:10:11 ; Search time 102.248 Seconds
(without alignment)

10230.248 Million cell updates/sec

Title: US-10-713-381-2_COPY_1155_1311

Perfect score: 157

Sequence: 1 cgtgcattctacatgcat.....cgcatatgcttcctgaac 157

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 7389322 seqs, 333185599 residues

Total number of hits satisfying chosen parameters: 14778644

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US10J_PUBCOMB.seq:*
22: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	157	100.0	255	20	US-10-713-381-9
2	157	100.0	1394	20	US-10-713-381-1
3	157	100.0	1394	20	US-10-713-381-2
4	146	93.0	158	20	US-10-713-381-3
5	50	31.8	50	20	US-10-713-381-5
6	40	25.5	40	20	US-10-713-381-6
7	32.2	20.5	399	19	US-10-674-124A-9081

C	8	31.6	20.1	2352	17	US-10-282-122A-36272	Sequence 36272, A
C	9	30.8	19.6	1122	19	US-10-437-963-1496	Sequence 1496, Ap
	10	30.6	19.5	559	13	US-10-027-632-244811	Sequence 244811,
	11	30.6	19.5	559	17	US-10-027-632-244811	Sequence 244811,
	12	30.6	19.5	79467	18	US-10-052-482-223	Sequence 223, App
	13	30.4	19.4	52001	19	US-10-317-273-11	Sequence 11, Appl
	14	30	19.1	30	20	US-10-713-381-4	Sequence 4, Appl
	15	30	19.1	843	13	US-10-027-632-170626	Sequence 170626,
	16	30	19.1	843	17	US-10-027-632-170626	Sequence 170626,
	17	29.8	19.0	236	9	US-09-783-590-10490	Sequence 10490, A
C	18	29.8	19.0	691	18	US-10-424-559-57018	Sequence 57018, A
	19	29.8	19.0	2735	13	US-10-027-632-254612	Sequence 254612,
	20	29.8	19.0	2735	17	US-10-027-632-254612	Sequence 254612,
	21	29.8	19.0	1223197	13	US-10-027-632-179264	Sequence 179264,
	22	29.8	18.9	1223197	17	US-10-027-632-179264	Sequence 179264,
C	23	29.6	18.9	680	19	US-10-767-701-2154	Sequence 2154, Ap
C	24	29.2	18.6	1475	17	US-10-321-802-11	Sequence 11, Appl
C	25	29.2	18.6	1830121	17	US-10-329-670-1	Sequence 1, Appl
C	26	29.2	18.6	1830121	20	US-10-158-865-1	Sequence 1, Appl
C	27	29.2	18.6	1830121	22	US-10-981-687-1	Sequence 1, Appl
C	28	29	18.5	358	20	US-10-425-115-178606	Sequence 178606, A
	29	29	18.5	553	13	US-10-027-632-222444	Sequence 222444,
	30	29	18.5	553	13	US-10-027-632-222444	Sequence 222444,
	31	29	18.5	553	17	US-10-027-632-222444	Sequence 222444,
C	32	29	18.5	553	17	US-10-027-632-222445	Sequence 222445,
C	33	29	18.5	7192	19	US-10-437-963-85365	Sequence 85365, A
	34	28.8	18.3	732	13	US-10-027-632-21721	Sequence 21721, A
	35	28.8	18.3	732	13	US-10-027-632-21722	Sequence 21722, A
	36	28.8	18.3	732	17	US-10-027-632-21722	Sequence 21721, A
	37	28.8	18.3	732	17	US-10-027-632-21722	Sequence 21722, A
	38	28.8	18.3	733	13	US-10-027-632-127179	Sequence 127179,
	39	28.8	18.3	733	13	US-10-027-632-127180	Sequence 127180,
	40	28.8	18.3	733	17	US-10-027-632-127180	Sequence 127180,
C	41	28.8	18.3	715517	13	US-10-027-632-53712	Sequence 53712, A
C	42	28.8	18.3	715517	17	US-10-027-632-53712	Sequence 53712, A
C	43	28.8	18.3	715517	20	US-10-425-115-141945	Sequence 141945,
	44	28.6	18.2	220	20	US-10-425-115-17846	Sequence 17846, A
	45	28.6	18.2	280	20	US-10-425-115-17846	Sequence 17846, A

ALIGNMENTS

RESULT 1
US-10-713-381-9
; Sequence 9, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentlin Ver. 2.0
; SEQ ID NO 9
; LENGTH: 255
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-9

Query Match 100.0%, Score 157, DB 20, Length 255,
Best Local Similarity 100.0%, Pred. No. 1.7e-43;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY 1 cgtgcattctacatgcatctactacatgcttttcacacgcttgcgttttcacatgcttc 60

```
Db      15 CGGTGCATCTCACATGCGATACATGCTGTGTTCAACCGTTGCTTGTTCATGCTCC 74
Qy      61 AAGCCTTGCTTATTTGGAACCAAGAGATACCTACTCCCAACATCATCTTACTCATG 120
Db      75 AAGCCTTGCTTATTTGGAACCAAGAGATACCTACTCCCAACATCATCTTACTCATG 134
Qy      121 CAACCTTCATGCAACGACATATGTTTCTGTAAC 157
Db      135 CAACCTTCATGCAACGACATATGTTTCTGTAAC 171

RESULT 2
US-10-713-381-1
; Sequence 1, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAAT, CARL W.
; APPLICANT: HUFEMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-1

Query Match      100.0%; Score 157; DB 20; Length 1394;
Best Local Similarity 100.0%; Pred. No. 3,1e-43;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CGGTGCATCTCACATGCGATACATGCTGTGTTCAACCGTTGCTTGTTCATGCTCC 60
Db      1155 CGGTGCATCTCACATGCGATACATGCTGTGTTCAACCGTTGCTTGTTCATGCTCC 1214
Qy      61 AAGCCTTGCTTATTTGGAACCAAGAGATACCTACTCCCAACATCATCTTACTCATG 120
Db      1215 AAGCCTTGCTTATTTGGAACCAAGAGATACCTACTCCCAACATCATCTTACTCATG 1274
Qy      121 CAACCTTCATGCAACGACATATGTTTCTGTAAC 157
Db      1275 CAACCTTCATGCAACGACATATGTTTCTGTAAC 1311

RESULT 3
US-10-713-381-2
; Sequence 2, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAAT, CARL W.
; APPLICANT: HUFEMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
```

```
; SEQ ID NO 2
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-2

Query Match      100.0%; Score 157; DB 20; Length 1394;
Best Local Similarity 100.0%; Pred. No. 3,1e-43;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CGGTGCATCTCACATGCGATACATGCTGTGTTCAACCGTTGCTTGTTCATGCTCC 60
Db      1155 CGGTGCATCTCACATGCGATACATGCTGTGTTCAACCGTTGCTTGTTCATGCTCC 1214
Qy      61 AAGCCTTGCTTATTTGGAACCAAGAGATACCTACTCCCAACATCATCTTACTCATG 120
Db      1215 AAGCCTTGCTTATTTGGAACCAAGAGATACCTACTCCCAACATCATCTTACTCATG 1274
Qy      121 CAACCTTCATGCAACGACATATGTTTCTGTAAC 157
Db      1275 CAACCTTCATGCAACGACATATGTTTCTGTAAC 1311

RESULT 4
US-10-713-381-3
; Sequence 3, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAAT, CARL W.
; APPLICANT: HUFEMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 158
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-3

Query Match      93.0%; Score 146; DB 20; Length 158;
Best Local Similarity 99.4%; Pred. No. 8,2e-40;
Matches 157; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy      1 CGGTGCATCTCACATGCGATACATGCTGTGTTCAACCGTTGCTTGTTCATGCTCC 59
Db      1 CGGTGCATCTCACATGCGATACATGCTGTGTTCAACCGTTGCTTGTTCATGCTCC 60
Qy      60 CAAGCTTGCTTATTTGGAACCAAGAGATACCTACTCCCAACATCATCTTACTCAT 119
Db      61 CAAGCTTGCTTATTTGGAACCAAGAGATACCTACTCCCAACATCATCTTACTCAT 120
Qy      120 GCAACTTCATGCAACGACATATGTTTCTGTAAC 157
Db      121 GCAACTTCATGCAACGACATATGTTTCTGTAAC 158

RESULT 5
US-10-713-381-5
; Sequence 5, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAAT, CARL W.
```

```
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent In Ver. 2.0
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-5

Query Match          31.8%; Score 50; DB 20; Length 50;
Best Local Similarity 100.0%; Pred. No. 4,6e-07;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      75 CTGAACCAAGAGATACCTACCTCCCAACATCCATCTTACTCATGCAC 124
DB      1 CTGAACCAAGAGATACCTACCTCCCAACATCCATCTTACTCATGCAC 50

RESULT 6
US-10-713-381-6
; Sequence 6, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent In Ver. 2.0
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-6

Query Match          25.5%; Score 40; DB 20; Length 40;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      85 AGGATACCTACTCTCCCAACATCCATCTTACTCATGCAC 124
DB      1 AGGATACCTACTCTCCCAACATCCATCTTACTCATGCAC 40

RESULT 7
US-10-674-124A-9081/c
; Sequence 9081, Application US/10674124A
; Publication No. US2004019797A1
; GENERAL INFORMATION:
; APPLICANT: INOKO, Hideoshl
; APPLICANT: TAMURA, Gen
; TITLE OF INVENTION: GENE MAPPING METHOD USING MICROSATELLITE
; FILE REFERENCE: ORIN-003CIP
; CURRENT APPLICATION NUMBER: US/10/674,124A
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 10/257,511
```

```
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: PCT/JP00/07621
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: JP2000-112699
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: JP2002-327516
; PRIOR FILING DATE: 2002-09-28
; PRIOR APPLICATION NUMBER: JP2002-383869
; PRIOR FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 27110
; SEQ ID NO 9081
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: chr5:fa.07f1r.139538206
; FEATURE:
; OTHER INFORMATION: Located on chromosome 5
; FEATURE:
; OTHER INFORMATION: Distance between a terminus base of telomere on
; OTHER INFORMATION: chromosomal short arm and 5'-terminus of this base
; FEATURE:
; OTHER INFORMATION: sequence : 128437212
; OTHER INFORMATION: Distance between 3'-terminus of neighbour sequence of
; OTHER INFORMATION: sequence listing upward to telomere on chromosomal short arm and
; OTHER INFORMATION: 5'-terminus of this base sequence : 250186
US-10-674-124A-9081

Query Match          20.5%; Score 32.2; DB 19; Length 399;
Best Local Similarity 61.2%; Pred. No. 1.3;
Matches 52; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY      72 ATTGACCAAGAGATACCTACCTCCCAACATCCATCTTACTCATGCAC 131
DB      239 ATTGGAAGAAAAAGTGGGCCACCTCAAAAAACCAACATACATCTTAATAAAA 180

QY      132 CAACACGACATATCTTCTCGAA 156
DB      179 AAAAAACCAACATTTGTATGAA 155

RESULT 8
US-10-282-122A-36272/c
; Sequence 36272, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangbu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
```

```
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 36272
/ LENGTH: 2352
/ TYPE: DNA
/ ORGANISM: Streptococcus mutans
US-10-282-122A-36272
```

```
Query Match      20.1%; Score 31.6; DB 17; Length 2352;
Best Local Similarity 58.5%; Pred. No. 3.9;
Matches 55; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
```

```
QY 62 AGCTTGCTTATTTCTGAACCAAGATACCTACCTCCCAACATCATCTTACTCATGC 121
    |||||
DB 234 ATCTGACTGTGCGCCAAAGCCAAAGAGATCATCTCTGAAAATCTGTGCTACATCTGC 175
    |||||
QY 122 AACTTCATGCAACACGCGCATATGTTTCTCTGA 155
    |||||
DB 174 AATTTGACGAAAGAAAGCTTCAATTTTTCCTGA 141
    |||||
```

RESULT 9

```
US-10-437-963-1496/C
/ Sequence 1496, Application US/10437963
/ Publication No. US20040123343A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Wu, Wei
/ APPLICANT: Boukharov, Andrey A.
/ APPLICANT: Barbazuk, Brad
/ APPLICANT: Li, Ping
/ TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53221)B
/ CURRENT APPLICATION NUMBER: US/10/437,963
/ NUMBER OF SEQ ID NOS: 204966
/ SEQ ID NO 1496
/ LENGTH: 1122
/ TYPE: DNA
/ ORGANISM: Oryza sativa
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT4530_101357C.1
US-10-437-963-1496
```

```
Query Match      19.6%; Score 30.8; DB 19; Length 1122;
Best Local Similarity 57.1%; Pred. No. 5.6;
Matches 56; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
```

```
QY 60 CAAGCTTGCTTATTTCTGAACCAAGATACCTACCTCCCAACATCATCTTACTCAT 119
    |||||
DB 313 CAGGCTCGCATATTTGCTATCCCGGATCATTAACTCAACAAAGCCATATTCATTCT 254
    |||||
QY 120 GCAACTTCATGCAACACGCGCATATGTTTCTCTGAC 157
    |||||
DB 253 TCAGATCGCAAAATCGACCTCCCATATCTGTCTGAAC 216
    |||||
```

```
RESULT 10
US-10-027-632-244811
```

```
/ Sequence 244811, Application US/10027632
/ Publication No. US20020198371A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ TITLE OF INVENTION: Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827,129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ CURRENT FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 244811
/ LENGTH: 559
/ TYPE: DNA
/ ORGANISM: Human
US-10-027-632-244811
```

```
Query Match      19.5%; Score 30.6; DB 13; Length 559;
Best Local Similarity 51.9%; Pred. No. 5.1;
Matches 69; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
```

```
QY 24 ACATGCTTGTCACCGTTGCTGTTGTCATGTCGAAGCCTTGCTATTTGACCAA 83
    |||||
DB 380 ACATGCAATTTTGAAGCTTTGCTGATTTGCTCATCTTGCGAGGTACTTTTCTAGATTCA 439
    |||||
QY 84 GAGGATACCTACTCCCAACACATCATCTTACTCATGCAACTTCATGCAACACGCGCA 143
    |||||
DB 440 TCTGCTCCCTTACTTTAAAGATTTCATGGAAGTCTTACAAATCCATAGCACACTGAAC 499
    |||||
QY 144 TATGTTCTCTGAA 156
    |||||
DB 500 AATTTCTTCCCA 512
    |||||
```

RESULT 11

```
US-10-027-632-244811
/ Sequence 244811, Application US/10027632
/ Publication No. US20030204075A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ TITLE OF INVENTION: Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827,129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ CURRENT FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
```

```

; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 244811
; LENGTH: 559
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-244811

Query Match          19.5%; Score 30.6; DB 17; Length 559;
Best Local Similarity 51.9%; Pred. No. 5.1;
Matches 69; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 24 ACATGCTGTTCACGCTGTGCTTCTTCATGCTGCAAGCCTTGCTATTCTGAACCA 83
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 380 ACATGCTGTTCACGCTGTGCTTCTTCATGCTGCAAGCCTTGCTATTCTGAATTC 439
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 84 GAGATTAACCTACTCCCAACATCATCTTACTCATGCACTTCATGCAACAGCACA 143
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 440 TCGCTCCCTACCTTTAAAGATTCCATGGAAGTCTCAAAATCCCATACACACTGAC 499
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 144 TATGTTTCTGAA 156
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 500 ATTTCTTCCCA 512
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 12
US-10-052-482-223
; Sequence 223, Application US/10052482
; Publication No. US2004072264A1
; GENERAL INFORMATION:
; APPLICANT: Engelhard, Eric
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: A-71087/RMS/DCF
; CURRENT APPLICATION NUMBER: US/10/052,482
; PRIOR FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 223
; LENGTH: 79467
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5502)..(5521)
; OTHER INFORMATION: "n" at positions 5502 to 5521 can be any base
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4095)..(4369)
; OTHER INFORMATION: "n" at positions 4099 to 4369 can be any base
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5502)..(5521)
; OTHER INFORMATION: "n" at positions 5502 to 5521 can be any base
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10089)..(10620)
; OTHER INFORMATION: "n" at positions 10089 to 10620 can be any base
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (13273)..(13370)
; OTHER INFORMATION: "n" at positions 13273 to 13370 can be any base
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (20762)..(20781)
; OTHER INFORMATION: "n" at positions 20762 to 20781 can be any base
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (30751)..(30916)
; OTHER INFORMATION: "n" at positions 30751 to 30916 can be any base
; NAME/KEY: misc feature
; LOCATION: (46579)..(46772)
```

```

; OTHER INFORMATION: "n" at positions 46759 to 46772 can be any base
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (77382)..(77401)
; OTHER INFORMATION: "n" at positions 77372 to 77401 can be any base
US-10-052-482-223

Query Match          19.5%; Score 30.6; DB 18; Length 79467;
Best Local Similarity 56.4%; Pred. No. 31;
Matches 57; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 51 TCCATGCTCAACGCTTGCTATTCTGCAACGAGATACCTACTCCCAACATTCAT 110
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 53779 TCCATGCTCAACGCTTGCTATTCTGCAACGAGATACCTACTCCCAACATTCAT 53838
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 111 CTACTCATGCAACTTCATGCAACGAGATATGTTTC 151
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 53839 CACACACACACACACACACACACACACACACACAGCTTGCTC 53879
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13
US-10-317-273-11
; Sequence 11, Application US/10317273
; Publication No. US20040110158A1
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Tamara Balac Sipes
; TITLE OF INVENTION: MODULATION OF MITOCHONDRIAL RIBOSOMAL PROTEIN L13 EXPRESSION
; FILE REFERENCE: RTS-0478
; CURRENT APPLICATION NUMBER: US/10/317,273
; CURRENT FILING DATE: 2002-12-10
; NUMBER OF SEQ ID NOS: 159
; SEQ ID NO 11
; LENGTH: 52001
; TYPE: DNA
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 11121, 11122, 11123, 11124, 11125, 11126, 11127, 11128, 11129,
; LOCATION: 11130, 11131, 11132, 11133, 11134, 11135, 11136, 11137,
; LOCATION: 11138, 11139, 11140, 11141, 11142, 11143, 11144, 11145,
; LOCATION: 11146, 11147, 11148, 11149, 11150, 11151, 11152, 11153
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 11154, 11155, 11156, 11157, 11158, 11159, 11160, 11161, 11162,
; LOCATION: 11163, 11164, 11165, 11166, 11167, 11168, 11169, 11170,
; LOCATION: 11171, 11172, 11173, 11174, 11175, 11176, 11177, 11178,
; LOCATION: 11179, 11180, 11181, 11182, 11183, 11184, 11185, 11186
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 11187, 11188, 11189, 11190, 11191, 11192, 11193, 11194, 11195,
; LOCATION: 11196, 11197, 11198, 11199, 11200, 11201, 11202, 11203,
; LOCATION: 11204, 11205, 11206, 11207, 11208, 11209, 11210, 11211,
; LOCATION: 11212, 11213, 11214, 11215, 11216, 11217, 11218, 11219
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 11220
; OTHER INFORMATION: n = A,T,C or G
US-10-317-273-11

Query Match          19.4%; Score 30.4; DB 19; Length 52001;
Best Local Similarity 61.2%; Pred. No. 31;
Matches 49; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 50 TTCATGCTCAACGCTTGCTATTCTGCAACGAGATACCTACTCCCAACATTCAT 109
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 20055 TTTCTTTTCTAAACTGAATTCATGCAAAAGAGATATCAATTTTAAACTACCA 20114
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 110 TCTTACTCATGCAACTTCA 129
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

DB 20115 TCTTCTCAGCAAACTTCTA 20134

RESULT 14
US-10-713-381-4

Sequence 4, Application US/10713381
Publication No. US20040221331A1
GENERAL INFORMATION:
APPLICANT: ALBERTSEN, MARC C.
APPLICANT: FOX, TIMOTHY W.
APPLICANT: GARNATT, CARL W.
APPLICANT: HUFEMAN, GARY
APPLICANT: KENDALL, TIMMY L.
TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
FILE REFERENCE: 578R
CURRENT APPLICATION NUMBER: US/10/713,381
CURRENT FILING DATE: 2003-11-14
PRIOR APPLICATION NUMBER: 08/880,499
PRIOR FILING DATE: 1997-06-23
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 30
TYPE: DNA
ORGANISM: Zea mays
US-10-713-381-4

Query Match 19.1%; Score 30; DB 20; Length 30;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 CATGCTGTTCACCGTTCGTTGTTCCCA 54
DB 1 CATGCTGTTCACCGTTCGTTGTTCCA 30

RESULT 15

US-10-027-632-170626
Sequence 170626, Application US/10027632
Publication No. US20020198371A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
TITLE OF INVENTION: Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 170626
LENGTH: 843
TYPE: DNA
ORGANISM: Human
US-10-027-632-170626

Query Match 19.1%; Score 30; DB 13; Length 843;
Best Local Similarity 51.5%; Pred. No. 9.5;
Matches 69; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 6 CATTCAGATGGCATATCATGTTTCAACCGTTCGTTGTTCCATGCCAAGCC 65
DB 552 CAGCACCCAGGCCAGCTCAAGCAACGTCAGCTCTCTGAGGGGCCACCATCCATGCC 611
QY 66 TTGCTTATTTCTGAACCAAGAGATACCTCCAAACATTCATCTTACTGATGCACT 125
DB 612 TTGCATGTGCTGGAAGATGGGGGCTTCGGCTCCCTGCTCATCATCAACGCTGGGAGGCCCTCC 671
QY 126 TCCATGCCAAACACG 139
DB 672 TGCATCCGAACACG 685

Search completed: September 15, 2005, 20:45:58
Job time : 106.248 secs

QY	61	AAGCTTGCCATTCTTGGAACCAAGGAGATACCTACTCCCAACATCATCTTACTCATG	120
DB	361	AAGCTTGCTATTCTTGGAACCAAGGAGATACCTACTCCCAACATCATCTTACTCATG	302
QY	121	CAACTTCATGCAAAACAGCAATATGTTTCTCGAAC	157
DB	301	CAACTTCATGCAAAACAGCAATATGTTTCTCGAAC	265
RESULT 2			
CG224225		915 bp	DNA
CG224225			linear
CG224225			GSS 22-AUG-2003
DEFINITION	OG1AG08TV.ZM.0.7.1.5_KB_Zea_mays_genomic_clone_ZM16B15,		
LOCUS	CG224225		
ACCESSION	CG224225		
VERSION	CG224225.1		
KEYWORDS	GSS.		
SOURCE	Zea_mays		
ORGANISM	Zea_mays		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.		
AUTHORS	1 (base 1 to 915) Whitelaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T., Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Cleek,R.W., Nunberg,A., Robbins,D. and Lakey,N.		
TITLE	Consortium for Maize Genomics		
JOURNAL	Unpublished (2002)		
COMMENT	Other_GSSs: OG1AG08TH Contact: Cathy Whitelaw TIGR		
FEATURES	712 Medical Center Drive, Rockville, MD 20850, USA Tel.: 301-838-5843 Fax: 301-838-0208 Email: whitelaw@cigr.org Seq primer: TP Class: sheared ends.		
SOURCE	Location/Qualifiers 1..915 /organism="Zea_mays" /mol_type="genomic DNA" /strain="B73" /db_xref="taxon:4577" /clone="ZM16B15" /clone_lib="ZM.0.7.1.5_KB" /note="Vector: pBESK-1_Site_1: HincII, 0.7-1.5 kb methylation filtered genomic DNA library"		
ORIGIN			
Query Match	100.0%; Score 157; DB 9; Length 915;		
Best Local Similarity	100.0%; Pred. No. 1.9e-39;		
Matches 157; Conservative	0; Mismatches 0; Indels 0; Gaps 0;		
QY	1	CGTGTGATCTGCATGAGATACATCATGCTTGTTCAACGGTTCGTCCTTGTTCATCGTCC	60
DB	534	CGTGTGATCTGCATGAGATACATCATGCTTGTTCAACGGTTCGTCCTTGTTCATCGTCC	593
QY	61	AAGCTTGCCATTCTTGGAACCAAGGAGATACCTACTCCCAACATCATCTTACTCATG	120
DB	594	AAGCTTGCCATTCTTGGAACCAAGGAGATACCTACTCCCAACATCATCTTACTCATG	653
QY	121	CAACTTCATGCAAAACAGCAATATGTTTCTCGAAC	157
DB	654	CAACTTCATGCAAAACAGCAATATGTTTCTCGAAC	690
RESULT 3			
CG656933		963 bp	DNA
CG656933			linear
CG656933			GSS 19-JUN-2003
DEFINITION	OGNDQ02TM.ZM.0.7.1.5_KB_Zea_mays_genomic_clone_ZM16B15,		
LOCUS	CG656933		
ACCESSION	CG656933		

VERSION	KEYWORDS	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
CC656933.1	GI:32060225	Zea mays	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoidae; Andropogoneae; Zea.	1 (bases 1 to 963)	WhiteLaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Reenick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nunez,A., Robbins,D. and Lakey,N.	Consortium for Maize Genomics Unpublished (2002)	Other GSSs: OGDQ020TV Contact: Cathy WhiteLaw
9712	Medical Center Drive, Rockville, MD 20850, USA	Tel: 301-838-5843 Fax: 301-838-0208 Email: whiteLaw@tigr.org Seq primer: TR Class: sheared ends.	Location/Qualifiers	1..963	/organism="Zea mays" /mol_type="genomic DNA" /strain="B73" /db_xref="taxon:4577" /clone_lib="ZMWBMA0554015" /clone_lib="ZM_0.7.1.5_KB" /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"		
Query Match	100.0%; Score 157; DB 9; Length 963;	Best Local Similarity	100.0%; Pred No. 2e-39;	Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	CGTGTCACTTCACATGGCATACTACATGCTTGTGTCACCGCTGCTTGTTCATGCTCC	60				
Db	752	CGTGTCACTTCACATGGCATACTACATGCTTGTGTCACCGCTGCTTGTTCATGCTCC	811				
Qy	61	AAGCTTGCCATTCTGGAACCAAGAGATATCCATCCCAACCAATCATCTTACATG	120				
Db	812	AAGCTTGCCATTCTGGAACCAAGAGATATCCATCCCAACCAATCATCTTACATG	871				
Qy	121	CAACTTCATGCAAAACGACACATATGTTCTCTGAAC	157				
Db	872	CAACTTCATGCAAAACGACACATATGTTCTCTGAAC	908				
RESULT 4							
CE588104/c	636 bp	DNA	linear	GSS 28-SEP-2003			
LOCUS	tigr-gss-dog-1700036359543	Dog Library	Canis familiaris genomic.				
DEFINITION							
ACCESSION	CE588104						
VERSION	CE588104.1	GI:36904885					
KEYWORDS	GSS.						
SOURCE	Canis familiaris (dog)						
ORGANISM	Canis familiaris						
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euteleia; Carnivora; Fissipedia; Canidae; Canis.						
REFERENCE	1 (bases 1 to 636)						
AUTHORS	Kirkneess,E.F., Batina,V., Halpern,A.L., Levy,S., Remington,K., Risch,D.B., Delcher,A.L., Pop,M., Wang,M., Fraser,C.M. and Venter,J.C.						
TITLE	The dog genome: survey sequencing and comparative analysis						
JOURNAL	Science 301 (5641), 1898-1903 (2003)						
MEDLINE	22875432						
PUBMED	14512627						
COMMENT	Contact: Kirkneess EF The Institute for Genomic Research						

LOCUS BB505306 700 bp mRNA linear EST 25-OCT-2001
DEFINITION BB505306 RIKEN full-length enriched, 10 days lactation, adult female mammary gland Mus musculus cDNA clone D730002G06, mRNA sequence.
ACCESSION BB505306
VERSION BB505306.2 GI:16442791
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 700)
AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Komu,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takehashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
Unpublished (2001)
On Jul 27, 2000 this sequence version replaced gi:5914268.
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenho-cho, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.jp, URL:http://genome-gsc.riken.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Komu,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
Riken integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Komu,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Komodo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamataka,I., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and Hayashizaki,Y.
Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome-gsc.riken.go.jp/) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

FEATURES

source

1..700
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="D730002G06"
/sex="female"
/tissue_type="mammary gland"
/dev_stage="10 days lactation, adult"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 10 days lactation, adult female mammary gland"
/note="Site_1: SalI; Site_2: BamHI, cDNA library was prepared and sequenced in Mouse Genome Encyclopedia

ORIGIN

Query Match 21.7%; Score 34; DB 2; Length 700;
Best Local Similarity 54.9%; Pred. No. 6.8;
Matches 67; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
QY 27 TCGTGTTCACACGTTGCTGTTCCATCGTCGAAGCCTTGCTATTCTGAACCAAG 86
DB 195 TTCTTTTTCCTCCCTCCCTCTGATTTCTTGACTCTTCACTTTTGCCCTTAAGT 254
QY 87 GATACCTACTCCCAACAATCCATCTTACTGCAACTTCATGCAACAGCACAAT 146
DB 255 GTTCTAGATTCCCAAGAAGCATTTTGACTAGAGTCAACATCACTGGAAT 314
QY 147 GT 148
DB 315 GT 316

RESULT 8
LOCUS CL026582 1677 bp DNA linear GSS 31-DEC-2003
DEFINITION CH216-23G24 Sp6.1 CH216 Xenopus tropicalis genomic clone
ACCESSION CL026582
VERSION CL026582.1 GI:40470443
KEYWORDS GSS.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae; Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 1677)
AUTHORS Kremitski,C., Carter,J., McPherson,J., Warren,W., Graves,T., Martis,E. and Wilson,R.
A physical map of the xenopus tropicalis genome
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 175000 Std Error: 0.00
Seq primer: Sp6 ATTGAGGACACATATAG
Class: BAC ends
High quality sequence start: 125
High quality sequence stop: 217.
Location/Qualifiers

FEATURES

source

1..1677
Location/Qualifiers
/organism="Xenopus tropicalis"
/mol_type="Genomic DNA"
/strain="Nigerian frog"
/db_xref="taxon:8364"
/clone="CH216-23G24"
/sex="male"
/cell_line="Stock 248 F7A2, inbred N7"
/clone_lib="CH216"
/note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis BAC library"

ORIGIN

Query Match 21.4%; Score 33.6; DB 9; Length 1677;
 Best Local Similarity 53.9%; Pred. No. 12;
 Matches 69; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 3 TGTGATCTCATGAGCACTTAAATGCTTGTTCACCGTTCGTTGTTCCATGTCACCA 62
 DB 769 TGTGATCTCATGAGCACTTAAATGCTTGTTCACCGTTCGTTGTTCCATGTCACCA 828
 QY 63 GCGTTCGCTTATTTGAAACCAAGAGATCTCTCCCAACATTCATTTACTATGCA 122
 DB 829 TCATCAACTATCTTCAAGTTCAATGTCACCTCTCCTCACTTCTTAATATTATTCAGACA 888
 QY 123 ACTTCAT 130
 DB 889 TCTTCAT 896

RESULT 9
 A1744861 440 bp mRNA linear EST 21-JUN-1999
 LOCUS cr16506.x1 NCI_CGAP_Ov23 Homo sapiens CDNA clone IMAGE:2218451 3,
 DEFINITION mRNA sequence.
 ACCESSION A1744861
 VERSION A1744861
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 440)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-rcmail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINL at:
 www.bio.lim.gov/bdrg/image/image.html
 Seq primer: -40UP from Gibco
 High quality sequence stop: 411.
 Location/Qualifiers
 1..440
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2218451"
 /issue_type="tumor, 5 pooled (see description)"
 /lab_host="DH10B"
 /clone_lib="NCI_CGAP_Ov23"
 /note="Organ: ovary; Vector: pCMV-SPORT6, Site 1: SalI,
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.35 kb. Tumor types include: mixed
 Mullerian tumor, papillary serous, clear cell, spindle
 cell. All are primary tumors, metastasis positive. Life
 Technologies catalog #: 11534-013"

ORIGIN
 Query Match 20.9%; Score 32.8; DB 1; Length 440;
 Best Local Similarity 56.5%; Pred. No. 14;
 Matches 61; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 46 CTTGTTTCATGTCGAAGCTTCTTATTTGAACCAAGATCTTCTCCCAACAA 105
 DB 259 CTTGTCACGCAAACTATATCAATATCTCTCCCAAGCAACACTCTTATTTT 318
 QY 106 TCCATCTTACTGACAACTTCATGCAACAGCACAATATGTTTCT 153

DB 319 TCCATCTTAAATTAACAGCGTCATGCTGCTTGAATATCTTCTCT 366

RESULT 10
 CR575123 804 bp mRNA linear EST 21-JUL-2004
 LOCUS CR575123 XGC-tailbud-head Xenopus tropicalis cDNA clone THDA024m23
 DEFINITION 5', mRNA sequence.
 ACCESSION CR575123
 VERSION CR575123
 KEYWORDS CR575123.1 GI:50462549
 SOURCE EST.
 ORGANISM Xenopus tropicalis (western clawed frog)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 Xenopodinae; Xenopus; Silurana.
 1 (bases 1 to 804)
 Sanger Xenopus tropicalis EST project 2001 (2004)
 CONTACT: Croning MDR
 UNPUBLISHED (2004)
 COMMENT Sanger Institute
 Hinxton, Cambridgeshire, CB10 1SA, UK
 Email: trop@sanger.ac.uk
 Sanger Xenopus tropicalis EST project 2001
 TROPICALIS_SEQUENCE_ID: THDA024m23.p1k5p6
 This sequence is from a Xenopus Gene Collection (XGC) library
 constructed by Nigel Garrett.
 Seq primer: SP6.
 Location/Qualifiers
 1..804
 /organism="Xenopus tropicalis"
 /mol_type="mRNA"
 /db_xref="taxon:8364"
 /clone="THDA024m23"
 /dev_stage="tailbud-head (stage 28-30)"
 /lab_host="Escherichia coli DH10B"
 /clone_lib="XGC-tailbud-head"
 /note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; CDNA
 was oligo dT primed from 5ug of poly A+ RNA from tailbud
 head. EcoRI-NotI cut CDNA was then ligated into pCS107
 with EcoRI at the 5' end and NotI at the 3' end."

ORIGIN
 Query Match 20.8%; Score 32.6; DB 7; Length 804;
 Best Local Similarity 63.3%; Pred. No. 20;
 Matches 50; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 74 TCTGAACCAAGAGATCTTCTCCCAACATCTTACTCATGCACTTCCATGCA 133
 DB 480 TCTGAATTTGAAAGGCCATCTCCCATGATTCATTTATGCAAGTATTAATTTT 539
 QY 134 AACAGCACAATATGTTTC 152
 DB 540 AAAAAGTATTTTTC 558

RESULT 11
 CNS003FV/c 1101 bp DNA linear GSS 03-JUN-1999
 LOCUS CNS003FV/c Drosophila melanogaster genome survey sequence ITB3 end of BAC #
 DEFINITION BACR08003 of RPL1 library from Drosophila melanogaster (fruit
 fly), genomic survey sequence.
 ACCESSION AL064657
 VERSION AL064657
 KEYWORDS AL064657.1 GI:4942009
 SOURCE GSS.
 ORGANISM Drosophila melanogaster (fruit fly)
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 1101)
 Genoscope.

JOURNAL

COMMENT	DIRECT Submission Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : segreif@genoscope.cns.fr) Web : www.genoscope.cns.fr)
ORIGIN	determination of this BAC-end and sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoko Osagawa and Aaron Mosmoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES	Location/Qualifiers 1..1101 /organism="Drosophila melanogaster" /mol_type="Genomic DNA" /db_xref="taxon:7227" /cclone="BACR08003" /clone_1lb="RPCI-98" /note="end : TET3"
SOURCE	
Query Match	20.8%; Score 32.6; DB 9; Length 1101;
Best Local Similarity	21.9%; Pred.No. 22;
Matches	Conservative 55; Mismatches 34; Indels 0; Gaps 0;
Oy	29 CTTGTTCACCGCTGCTTGGTGATCATGCAGCCTGCCATTCTGAACAAGA GA 88 : ::::: ::::: ::::: ::::: ::::: ::::: ::::: Db CBWTKDAAYBKDATKBSYCCMAAVSCCARDNYMACHDKCYCMCHWHNAHAA 938 997
Oy	89 TACTTA CTCCAACATCACTTACTA TGCAA GTTCATG CAACACGCAC 142 ::: :: :: :: :: :: :: :: :: :: :: :: :: Db WDNBKKBCCBAMAABRCRKDKKTYYMKTXBYCVCKAKGRSYKSICC 884 937
RESULT 12	BEG90384 347 bp mRNA linear EST 11-SEP-2000
LOCUS	uw66fi2.yl Soares_mammary_gland_MMLMG Mus musculus cDNA clone
DEFINITION	IMAGEJ3466991 5' , mRNA sequence.
ACCESION	BEG90384
VERSION	BEG90384.1 GI:10078008
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (baes 1 to 347) NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap. National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997) Contact: Robert Strauszberg, Ph.D. Email: cgepbx-re@mail.nih.gov This clone is available royalty-free through LNL / contact the IMAGE Consortium (info@image.llnl.gov) for further information. MGJ:1386351
REFERENCE	Seq primer:- 40RP from Gibco
AUTHORS	High quality sequence stop: 270.
TITLE	Location/Qualifers 1..347
JOURNAL COMMENT	/organism="Mus musculus" /mol_type="mRNA" /db_xref="taxon:10090" /cclone="IMAGEj3466991" /sex="female (lactating)"
FEATURES	
SOURCE	

ORIGIN	/tissue_type="mammary gland"	
	/lab_host="DHL0B"	
	/clone_idb="Soares_mammary_gland_NMLNG"	
	/note="Vector: pT73D-Pac (Pharmacia) with a modified	
	polylinker; 1st strand cDNA was prepared from mammary	
	gland tissue from a lactating female, and was then primed	
	with a Not I - Oligo(dT) primer. Double-stranded cDNA was	
	ligated to Eco RI adaptors (Pharmacia), digested with Not	
	I and cloned into the Not I and Eco RI sites of the	
	modified pT73 vector. Library is normalized. Library	
	was constructed by Benito Soares and M. Fatima Bonaldo."	
Query Match	20.6%;	Score 32.4; DB 2; Length 347;
Best Local Similarity	54.1%;	Pred. No. 18;
Matches	66; Conservative	0; Mismatches 56; Indels 0; Gaps 0;
QY	27	TGCTTTTTCACCGTTGCTCTTTGTCACATGTCACAGCCTTGCTATTTCGACCAAG 86
DB	148	TTCTTTTTCACCTCCCTCTGATTTTGTGTACTGCTTCACTTTTGCGCTTAAATT 207
QY	87	GATACCTACTCCCAACATCATCTACTCATGCACTTCACGCAACAGCACATAT 146
DB	208	GTTCTAGATTCCCAAGAGCATTGTGGACTACGACGCAACATCTACACCTGAAT 267
QY	147 GT 148	
DB	268 GT 269	
RESULT 13	AJ746802/c 483 bp mRNA linear EST 07-JUL-2004	
LOCUS	AJ746802 forward - stimulated minus unstimulated macrophage Sue	
DEFINITION	scrofa cDNA clone F_C0001825C_D07, mRNA sequence.	
ACCESSION	AJ746802	
VERSION	AJ746802.1 GI:49916882	
KEYWORDS	EST.	
SOURCE	Sus scrofa (pig)	
ORGANISM	Sus scrofa	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.	
REFERENCE	1 (bases 1 to 483)	
AUTHORS	Hopwood,P.A., Zhang,F., Lowden,S., Talbot,R., Burt,D., Archibald,A.	
	and Dixon,L.	
TITLE	Development of a porcine cDNA microarray	
JOURNAL	Unpublished (2004)	
COMMENT	Contact: Hopwood PA	
	Dept. of Preclinical Veterinary Sciences	
	Royal School for Veterinary Studies	
	Summerhall, Edinburgh, EH9 1QH, UNITED KINGDOM	
	Sequencing was performed by ARK genomics. This clone is available	
	from ARK- Genomics, Roslin Institute, Roslin, Midlothian EH25 9PS,	
	UK. See www.ark-genomics.org or contact info@arkgenomics.org.	
FEATURES	Location/Qualifiers	
source	1..483	
	/organism="Sus scrofa"	
	/mol_type="mRNA"	
	/db_xref="taxon:9823"	
	/clone="F_C0001825C_D07"	
	/tissue_type="lung"	
	/cell_type="macrophage"	
	/clone_idb="forward - stimulated minus unstimulated	
	macrophage"	
ORIGIN		
Query Match	20.6%;	Score 32.4; DB 1; Length 483;
Best Local Similarity	55.3%;	Pred. No. 20;
Matches	63; Conservative	0; Mismatches 51; Indels 0; Gaps 0;
QY	21	ACTACATGCTTTTCACACCGTTGCTCTTTGTCACATGTCACAGCCTTGCTATTCTGAC 80
DB	371	ACCAATCTTGAATGAATGAAGCTTTTTTTTTTTTAAAAAAGCCCAATATATCTTAAC 312

QY 81 CAGAGATACCTACTCCCAACATTCATCTTACTCATGCAACTTCCATGCA 134
DB 311 ACAGTAAAGTGCACCTATCATCTTCAATTAATTAAGATTTTCACAGGAA 258

RESULT 14
AM990724/c 497 bp mRNA linear EST 02-JUN-2000
LOCUS u110b11.x1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone
DEFINITION IMAGE:1510941 3', mRNA sequence.
ACCESSION AM990724
VERSION AM990724.1 GI:8186451
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
COMMENT Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
Tumor Gene Index
Unpublished (1997)
JOURNAL Other ESTs: u110b11.y1
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapsb@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:937793
High quality sequence stop: 459.
location/Qualifiers
1..497
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:1510941"
/sex="female (lactating)"
/tissue_type="mammary gland"
/lab_host="DH10B"
/clone_lib="Soares_mammary_gland_NMLMG"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from mammary
gland tissue from a lactating female, and was then primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pT73 vector. Library is normalized. Library
was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN
Query Match 20.6%; Score 32.4; DB 2; Length 497;
Best Local Similarity 54.1%; Pred. No. 20;
Matches 66; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 27 TCGTTGTTCAACCGTTCGTTGTCATGTCGCAAGCTTGCTATTGCAACCAAG 86
DB 490 TTTCTTTTTCACCTCCCTCTGATTTCTTTGTACTTCATCTTATTTGGCCTTTAAGTT 431

QY 87 GATACCTACTCCCAACATTCATCTTACTCATGCAACTTCAGCAACAGCAGACATAT 146
DB 430 GTTCTAGATTCCCAAGAGCCATTGTCACCTACGACGCAACATTAACACCTGAAAT 371

QY 147 GT 148
DB 370 GT 369

RESULT 15
BE632212 498 bp mRNA linear EST 25-AUG-2000
LOCUS u111f01.x1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone
DEFINITION IMAGE:3371641 3', mRNA sequence.
ACCESSION BE632212

VERSION BE632212.1 GI:9914830
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
COMMENT Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
Tumor Gene Index
Unpublished (1997)
JOURNAL Other ESTs: u111f01.y2
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapsb@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:1081245
High quality sequence stop: 459.
location/Qualifiers
1..498
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:3371641"
/sex="female (lactating)"
/tissue_type="mammary gland"
/lab_host="DH10B"
/clone_lib="Soares_mammary_gland_NMLMG"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from mammary
gland tissue from a lactating female, and was then primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pT73 vector. Library is normalized. Library
was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN
Query Match 20.6%; Score 32.4; DB 2; Length 498;
Best Local Similarity 54.1%; Pred. No. 20;
Matches 66; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 27 TCGTTGTTCAACCGTTCGTTGTCATGTCGCAAGCTTGCTATTGCAACCAAG 86
DB 491 TTTCTTTTTCACCTCCCTCTGATTTCTTTGTACTTCATCTTATTTGGCCTTTAAGTT 432

QY 87 GATACCTACTCCCAACATTCATCTTACTCATGCAACTTCAGCAACAGCAGACATAT 146
DB 431 GTTCTAGATTCCCAAGAGCCATTGTCACCTACGACGCAACATTAACACCTGAAAT 372

QY 147 GT 148
DB 371 GT 370

Search completed: September 16, 2005, 08:08:43
Job time : 490.661 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 08:25:31 ; Search time 100.667 Seconds
(without alignment)
14440.280 Million cell updates/sec

Title: US-10-713-381-2_COPY_1179_1208

Perfect score: 30
Sequence: 1 catgctgttcacacgcttcgtctgttcca 30

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmb1:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_srb:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	30	100.0	30	6	AX224397 Sequence
2	30	100.0	255	6	AX224402 Sequence
3	30	100.0	1394	6	AX224394 Sequence
4	30	100.0	1394	6	AX224395 Sequence
5	30	100.0	1394	6	BD062176 Male t1ss
6	30	100.0	1394	6	BD062177 Male t1ss
7	30	100.0	3343	8	AF360356 Zea mays
8	24.2	80.7	158	6	AX224396 Sequence
9	22	73.3	1304	6	AX523502 Sequence
10	21.6	72.0	187192	2	AC124638 Mus muscu
11	21.6	72.0	188427	2	AC151282 Mus muscu
12	21.2	70.7	129778	8	AC091123 Oryza sat
13	21.2	70.7	140729	8	AP006548 Oryza sat
14	21.2	70.7	153675	2	AP005710 Oryza sat
15	21.2	70.7	153146	8	AP004118 Oryza sat
16	21.2	70.7	158374	8	AP004786 Oryza sat
17	21.2	70.7	172058	8	AP003682 Oryza sat
18	21.2	70.7	178158	8	AC084320 Oryza sat
19	21	70.0	91298	2	AC119671 Oryza sat

C	20	21	70.0	93342	9	AC087245 Homo sapi
C	21	21	70.0	123472	8	AC120984 Oryza sat
C	22	21	70.0	142015	8	AC120884 Oryza sat
C	23	21	70.0	143681	8	AC082645 Oryza sat
C	24	21	70.0	156394	2	AC026558 Homo sapi
C	25	21	70.0	158839	8	AC090882 Oryza sat
C	26	21	70.0	162198	8	AP005916 Oryza sat
C	27	21	70.0	177790	8	AP005930 Oryza sat
C	28	21	70.0	187707	8	AC077693 Oryza sat
C	29	21	70.0	188654	10	AL626770 Mouse DNA
C	30	21	70.0	253492	2	AC102160 Mus muscu
C	31	21	70.0	300029	8	AE017121 Oryza sat
C	32	20.6	68.7	173088	8	AC109365 Oryza sat
C	33	20.6	68.7	196303	10	AC101349 Mus muscu
C	34	20.6	68.7	206372	10	EX005189 Mouse DNA
C	35	20.4	68.0	1026	14	H1M42890 Human tmm
C	36	20.4	68.0	1542	6	AX506929 Sequence
C	37	20.4	68.0	1762	8	AY049275 Arabidops
C	38	20.4	68.0	1682	8	AT525213 Arabidops
C	39	20.4	68.0	1913	8	AF242307 Euphorbia
C	40	20.4	68.0	1985	8	AC067922 Homo sapi
C	41	20.4	68.0	42868	9	AC115980 Mus muscu
C	42	20.4	68.0	65406	2	AC021665 Arabidops
C	43	20.4	68.0	80413	8	OSJN00036 Oryza sat
C	44	20.4	68.0	114051	8	AF06598 Oryza sat
C	45	20.4	68.0	146776	8	AP007223 Oryza sat

ALIGNMENTS

RESULT 1	AX224397	Sequence 4 from Patent WO0160997.	30 bp	DNA	linear	PAT 10-SEP-2001
LOCUS	AX224397					
DEFINITION	AX224397					
ACCESSION	AX224397.1	GI:15554639				
VERSION						
KEYWORDS						
SOURCE						
ORGANISM						
Zeae mays						
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.						
REFERENCE						
AUTHORS	Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.					
TITLE	Male tissue-preferred regulatory region and method of using same					
JOURNAL	Patent: WO 0160997-A 4 23-AUG-2001;					
PIONEER	HI-BRED INTERNATIONAL, INC. (US)					
FEATURES						
source	1..30	Location/Qualifiers				
	/organism="Zea mays"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:4577"					
ORIGIN						
Query Match	100.0%;	Score 30;	DB 6;	Length 30;		
Best Local Similarity	100.0%;	Pred. No. 0.03;				
Matches	30;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
Qy	1	CATGCTGTTCAACCGTTCGCTGTTCCA	30			
Db	1	CATGCTGTTCAACCGTTCGCTGTTCCA	30			
RESULT 2	AX224402	Sequence 9 from Patent WO0160997.	255 bp	DNA	linear	PAT 10-SEP-2001
DEFINITION	AX224402					
ACCESSION	AX224402.1	GI:15554644				
VERSION						
KEYWORDS						
SOURCE						
ORGANISM						
Zeae mays						

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
1
Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
Patent: WO 0160997-A 9 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)

FEATURES
source
1..255
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN

Query Match 100.0%; Score 30; DB 6; Length 255;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY
1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
|||||
39 CATGCTTGTTCAACCGTTCGTTGTTCCA 68

Db
39 CATGCTTGTTCAACCGTTCGTTGTTCCA 68

RESULT 3
AX224394 1394 bp DNA linear PAT 10-SRP-2001
LOCUS
DEFINITION
Sequence 1 from Patent WO0160997.
AX224394
VERSION
AX224394.1 GI:15554636
KEYWORDS
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
1
Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
Patent: WO 0160997-A 1 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)

FEATURES
source
1..1394
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN

Query Match 100.0%; Score 30; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY
1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
|||||
1179 CATGCTTGTTCAACCGTTCGTTGTTCCA 1208

Db
1179 CATGCTTGTTCAACCGTTCGTTGTTCCA 1208

RESULT 4
AX224395 1394 bp DNA linear PAT 10-SRP-2001
LOCUS
DEFINITION
Sequence 2 from Patent WO0160997.
AX224395
VERSION
AX224395.1 GI:15554637
KEYWORDS
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
1
Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
Patent: WO 0160997-A 2 23-AUG-2001;

PIONEER HI-BRED INTERNATIONAL, INC. (US)

FEATURES
source
1..1394
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN

Query Match 100.0%; Score 30; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY
1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
|||||
1179 CATGCTTGTTCAACCGTTCGTTGTTCCA 1208

Db
1179 CATGCTTGTTCAACCGTTCGTTGTTCCA 1208

RESULT 5
BD062176 1394 bp DNA linear PAT 27-AUG-2002
LOCUS
DEFINITION
Male tissue-preferred regulatory region and method of using same.
BD062176
VERSION
BD062176.1 GI:22607781
KEYWORDS
JP 2001520523-A/1.
Homo sapiens (human)
SOURCE
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
1
bases 1 to 1394
Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G.A. and Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
Patent: JP 2001520523-A 1 30-OCT-2001;
PIONEER HI BRED INTERNATIONAL INC

COMMENT
PN JP 2001520523-A/1
PD 30-OCT-2001
PF 19-JUN-1998 JP 1999504910
PR 23-JUN-1997 US 08/880499
PI MARC C ALBERTSEN,TIMOTHY W FOX,CARL W GARNAAAT,GARY A HUFFMAN,
PI TIMMY L KENDALL
PC C12N15/62,C12N15/29,C12N9/24,C12N9/22,C12N9/10,C12N9/00 PC
,C07K4/34,C12Q1/68,
PC A01H5/00
CC Strandedness: Single;
CC Topology: Linear;
FH key Location/Qualifiers.

FEATURES
source
1..1394
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 30; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY
1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
|||||
1179 CATGCTTGTTCAACCGTTCGTTGTTCCA 1208

Db
1179 CATGCTTGTTCAACCGTTCGTTGTTCCA 1208

RESULT 6
BD062177 1394 bp DNA linear PAT 27-AUG-2002
LOCUS
DEFINITION
Male tissue-preferred regulatory region and method of using same.
BD062177
VERSION
BD062177.1 GI:22607782
KEYWORDS
JP 2001520523-A/2.
Homo sapiens (human)
SOURCE
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 1394)
AUTHORS Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G.A. and Kendall,T.L.
TITLE Male tissue-preferred regulatory region and method of using same
JOURNAL Patent: JP 2001520523-A 2 30-OCT-2001;
PIONEER HI-BRED INTERNATIONAL INC
COMMENT PN JP 2001520523-A/2
PD 30-OCT-2001
PR 19-JUN-1998 JP 1999504910
PR 23-JUN-1997 US 08/880499
PI MARC C ALBERTSEN, TIMOTHY W FOX, CARL W GARNAA, GARY A HUFFMAN,
PI TIMMY L KENDALL
PC C12N15/82, C12N15/29, C12N9/24, C12N9/22, C12N9/10, C12N9/00 PC
C07K14/34, C12Q1/68,
PC A01H5/00
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers.
FEATURES
source 1..1394
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 30; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CATGCTTGTTCACCGTTCGTCCTTGTTC 30
Db 1179 CATGCTTGTTCACCGTTCGTCCTTGTTC 1208
RESULT 7
LOCUS AF360356 3343 bp DNA linear PLN 12-MAY-2001
DEFINITION Zea mays male fertility protein (M45) gene, complete cds.
ACCESSION AF360356
VERSION AF360356.1 GI:14028756
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 3343)
Fox,T.W., Trimnell,M.R. and Albertsen,M.C.
Cloning of M45, a gene required for male fertility from Zea mays
Unpublished
2 (bases 1 to 3343)
Fox,T.W., Trimnell,M.R. and Albertsen,M.C.
Direct Submission
Submitted (13-MAR-2001) Trait and Technology Development, Pioneer
Hi-Bred Intl. Inc., 7300 N.W. 62nd Ave., P.O. Box 1004, Johnston,
IA 50131-1004, USA
Location/Qualifiers
1..3343
/organism="Zea mays"
/mol_type="genomic DNA"
/culivars="B73"
/db_xref="taxon:4577"
/chromosome="9L"
<1392..>2942
/gene="M45"
join(<1392..1768,1898..2182,2280..2447,2534..>2942)
/gene="M45"
/product="male fertility protein"
join(1392..1768,1898..2182,2280..2447,2534..2942)
/gene="M45"
/codon_start=1
/product="male fertility protein"
/protein_id="AKS2489.1"

/db_xref="GI:14028757"
/translation="MKRULQRRRGDGIYQYPHLFFALALALVADPFGSLAE
DYRPVHLELAPGEVWGSMPDRNASRLRGRLEFGEVGPESIEFDLQGRPYGLA
DQVVRHMEBEAGWETFAVMNPDWSEVCANVSTTRQHKESFCGRPLGRFGE
TSELVYADAYYGLMVGQSGVASSVAREADSDPIRFANDLVHRNGSVFTDSMR
SKRDHLNITLIEGGTGRRLRYDPETSGVIVLKGIVFPNGVOISDHQFLPSETNG
RIMRWLEGPBAGVEVFPANLPGFPDNPVNSNGCGPWWAIDCCRTPAQSVFAKRPRLR
TIYKRPPLSLKVLIVMKARMTVTVALLDSEGRVVEVLDRGHGVKLVSEVREGRK
LWIGTVAHNHVIATIPYPLED"
ORIGIN
Query Match 100.0%; Score 30; DB 8; Length 3343;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CATGCTTGTTCACCGTTCGTCCTTGTTC 30
Db 1179 CATGCTTGTTCACCGTTCGTCCTTGTTC 1208
RESULT 8
LOCUS AX224396 158 bp DNA linear PAT 10-SEP-2001
DEFINITION Sequence 3 from Patent WO0160997.
ACCESSION AX224396
VERSION AX224396.1 GI:15554638
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1
Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
Patent: WO 0160997-A 3 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
Location/Qualifiers
1..158
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"
ORIGIN
Query Match 80.7%; Score 24.2; DB 6; Length 158;
Best Local Similarity 89.7%; Pred. No. 11;
Matches 26; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CATGCTTGTTCACCGTTCGTCCTTGTTC 29
Db 25 CATGCTTGTTCACCGTTCGTCCTTGTTC 53
RESULT 9
LOCUS AX523502 1304 bp DNA linear PAT 24-OCT-2002
DEFINITION Sequence 90 from Patent WO02064788.
ACCESSION AX523502
VERSION AX523502.1 GI:24412398
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1
Macina,R.A., Recipon,H., Chen,S.Y., Sun,Y. and Liu,C.
Compositions and methods relating to lung specific genes and
proteins
Patent: WO 02064788-A 90 22-AUG-2002;
Diadexus, Inc. (US)
Location/Qualifiers
1..1304
/organism="Homo sapiens"

ORIGIN
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

Query Match 73.3% Score 22; DB 6; Length 1304;
Best Local Similarity 83.3% Pred. No. 97;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
1 CATGCTTGTTCAACCGTTCGTTGTTCC 30
1155 CTTCTGTTCCCGCCGTTCTCTGTTCCA 1126

RESULT 10 AC124638 187192 bp DNA linear HTG 05-MAR-2004
AC124638/c MUS MUSCULUS CHROMOSOME 17 CLONE RP23-461H1 map 17, *** SEQUENCING
LOCUS IN PROGRESS ***; 10 unordered pieces.

ACCESSION AC124638 GI:45120281
VERSION AC124638.3
KEYWORDS HTG; HTGS PHASE1; HTGS_FULFILLTOP; HTGS_ACTIVEFIN.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE AUTHORS Birren,B., Nussbaum,C. and Lander,E.
TITLE Mus musculus chromosome 17, clone RP23-461H1
JOURNAL 1 (bases 1 to 187192)
REFERENCE 2 (bases 1 to 187192)
AUTHORS Birren,B., Linton,L., Nussbaum,C., Lander,E., Ali,A., Allen,N.,
JOURNAL Anderson,S., Barina,N., Bastien,V., Bloom,T., Boguslavsky,L.,
AUTHORS Boukhalil,B., Brown,A., Camarata,J., Campoliano,A., Chang,D.,
REFERENCE Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
AUTHORS Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
JOURNAL Faro,S., Ferreira,P., Fitzgerald,M., Fitzhugh,W., Gage,D.,
REFERENCE Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L.,
AUTHORS Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I.,
REFERENCE Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Larocque,K.,
AUTHORS Lamazares,R., Landers,T., Lehocsky,J., Levine,R., Lindblad-Toh,K.,
JOURNAL Liu,G., Maclean,C., Macdonald,P., Major,U., Margis,N.,
REFERENCE Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrum,J.,
AUTHORS Meneus,L., Mhova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
JOURNAL Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P.,
REFERENCE O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
AUTHORS Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C.,
JOURNAL Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S.,
REFERENCE Schupbach,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N.,
AUTHORS Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S.,
JOURNAL Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J.,
REFERENCE Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
AUTHORS Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE Direct Submission
JOURNAL Submitted (15-JUN-2002) Whitehead Institute/MIT Center for Genome
AUTHORS Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 187192)

Birren,B., Nussbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Arachchi,H.M., Barina,N., Bastien,V., Bloom,T.,
Boguslavsky,L., Boukhalil,B., Camarata,J., Chang,J., Choepel,Y.,
Collymore,A., Cook,A., Cooke,P., Corum,B., Dearellano,K.,
Diaz,J.S., Dodge,S., Doolley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., Maclean,C.,
Macdonald,P., Major,U., Manning,J., Matthews,C., McCarthy,M.,
Meldrum,J., Meneus,L., Mhova,T., Mlenga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schupbach,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,

TITLE Talamae,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
JOURNAL Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
AUTHORS Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

COMMENT Direct Submission
Submitted (05-MAR-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 5, 2004 this sequence version replaced GI:28893670.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green,P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute / MIT Center for Genome Research
Center code: WIRB
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
Project Information
Center project name: L24115
Center clone name: 461_H_1

* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 9926: contig of 9926 bp in length
* 9927 10026: gap of 100 bp
* 10027 20081: contig of 10055 bp in length
* 20082 20181: gap of 100 bp
* 20182 43259: contig of 23078 bp in length
* 43260 43359: gap of 100 bp
* 43360 45947: contig of 2588 bp in length
* 45948 46047: gap of 100 bp
* 46048 100940: contig of 54893 bp in length
* 100941 101040: gap of 100 bp
* 101041 106719: contig of 5679 bp in length
* 106720 106819: gap of 100 bp
* 106820 140370: contig of 33551 bp in length
* 140371 140470: gap of 100 bp
* 140471 147705: contig of 7235 bp in length
* 147706 147805: gap of 100 bp
* 147806 152565: contig of 4760 bp in length
* 152566 152665: gap of 100 bp
* 152666 187192: contig of 34527 bp in length.

FEATURES
source location/Qualifiers
1..187192
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="17"
/map="17"
/clone="RP23-461H1"
/clone_1fb="RP23-461H1"
/clone_1fb="RP23-461H1"

ORIGIN
Query Match 72.0% Score 21.6; DB 2; Length 187192;
Best Local Similarity 85.7% Pred. No. 1.4e+02;
Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 ATGCTTGTTCAACCGTTCGTTGTTCC 29
DB 52881 ATGCTTGTTCAACCGTTCGTTGTTCC 52854

RESULT 11 AC151282 188427 bp DNA linear HTG 24-AUG-2004
AC151282/c MUS MUSCULUS CHROMOSOME 17 CLONE RP23-461H1, *** SEQUENCING IN
LOCUS PROGRESS ***; 12 unordered pieces.
ACCESSION AC151282 GI:51511039
VERSION AC151282.1

```

KEYWORDS      HTG; HTGS_PHASE1.
SOURCE        Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE     1 (bases 1 to 188427)
AUTHORS       Wilson,R.K.
TITLE         The sequence of Mus musculus clone
JOURNAL       Unpublished
REFERENCE     2 (bases 1 to 188427)
AUTHORS       Wilson,R.K.
TITLE         Direct Submission
JOURNAL       Submitted (24-AUG-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA

COMMENT
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site:http://genome.wustl.edu
----- Project Information -----
Center project name: M_BA0461H01
----- Summary Statistics -----
Sequencing vector: M13; 0%
Sequencing vector: plasmid; 100%
Chemistry: Dye-terminator Big Dye; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 182668 bases at least Q40
Consensus quality: 183763 bases at least Q30
Consensus quality: 184341 bases at least Q20
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1      1129: contig of 1129 bp in length
*      1130      1229: gap of unknown length
*      1230      2832: contig of 1603 bp in length
*      2833      2932: gap of unknown length
*      2933      4658: contig of 1726 bp in length
*      4659      4759: gap of unknown length
*      4759      6427: contig of 1669 bp in length
*      6428      6527: gap of unknown length
*      6528      11124: contig of 4597 bp in length
*      11125      11224: gap of unknown length
*      11225      21177: contig of 9953 bp in length
*      21178      21277: gap of unknown length
*      21278      31319: contig of 10042 bp in length
*      31320      31419: gap of unknown length
*      31420      54758: contig of 23339 bp in length
*      54759      54858: gap of unknown length
*      54859      95804: contig of 40946 bp in length
*      95805      95904: gap of unknown length
*      95905      135186: contig of 39282 bp in length
*      135187      135286: gap of unknown length
*      135287      186683: contig of 5197 bp in length
*      186684      186783: gap of unknown length
*      186784      188427: contig of 1644 bp in length.
*
* Location/Qualifiers
*   .1.188427
*     /organism="Mus musculus"
*     /mol_type="genomic DNA"
*     /db_xref="taxon:10090"
*     /chromosome="17"
*     /clone="RP23-461H1"
*     .1..1129
*     /note="assembly_contig11"
*     misc_feature
*       1230..2832

```

	misc_feature	/note="assembly_name:Contig15"
	misc_feature	.933. .4658 /note="assembly_name:Contig16"
	misc_feature	.4759. .6427 /note="assembly_name:Contig19"
	misc_feature	.6528. .11124 /note="assembly_name:Contig20"
	misc_feature	.11225. .21177 /note="assembly_name:Contig21"
	misc_feature	.21278. .31319 /note="assembly_name:Contig22"
	misc_feature	.31420. .54758 /note="assembly_name:Contig23"
	misc_feature	.54859. .95804 /note="assembly_name:Contig24"
	misc_feature	.95905. .135186 /note="assembly_name:Contig25"
	misc_feature	.135287. .186693 /note="assembly_name:Contig26"
	misc_feature	.186784. .188427 /note="assembly_name:Contig6"
ORIGIN		
Query Match	72.0%; Score 21.6; DB 2;	Length 188427;
Best Local Similarity	85.7%; Pred. No. 1.4e+02;	
Matches 24; Conservative 0; Mismatches 4;	Indels 0; Gaps 0;	
Oy	2 ATGCTGTTCACCGTCCGGTTGTTGC 29	
Dd	 182832 ATGCTTGTAACAAGTAGATCATTCTGCC 182859	
RESULT 12		
AC091123/c	129778 bp DNA linear	PLN 09-AUG-2001
LOCUS	Oryza sativa chromosome 3 BAC OSUNBD0093EJ3 genomic sequence,	
DEFINITION	complete sequence.	
ACCESSION	AC091123	
VERSION	AC091123.4 GI:14670090	
KEYWORDS	HMG.	
SOURCE	Oryza sativa (japonica cultivar-group)	
ORGANISM	Oryza sativa (japonica cultivar-group) Euariyote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae; Euhartioideae; Oryzaceae; Oryza. 1 (bases 1 to 129778) Buell,C.R., Yuan,Q., Ouyang,S., Moffat,K.S., Hill,J.N., Gansberger,K., Brenner,M., Burgess,S., Hance,M., Shvartsbeyn,M., Telstrin,T., Riggs,F., Hsiiao,Y., Ziemann,V., Blunt,S., Pal,G., Vanaberg,S.E., Utecherack,T.R., Feldblum,T.V., Quackenbush,D., Salzburg,S.L., White,O. and Frazer,C.M. Oryza sativa chromosome 3 BAC OSUNBD0093EJ3 genomic sequence Unpublished 2 (bases 1 to 129778) Buell,R. Direct Submission Submitted (31-MAR-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA 3 (bases 1 to 129778) Buell,R. Direct Submission Submitted (11-JUL-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA 4 (bases 1 to 129778) Buell,R. Direct Submission Submitted (14-JUL-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA, rbuelletigr.org 5 (bases 1 to 129778) Buell,R. Direct Submission Submitted (18-JUL-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA, rbuelletigr.org	
TITLE	JOURNAL	
REFERENCE	AUTHORS	
TITLE	JOURNAL	
REFERENCE	AUTHORS	
TITLE	JOURNAL	
REFERENCE	AUTHORS	
TITLE	JOURNAL	

REFERENCE 6 (bases 1 to 129778)
AUTHORS Buell, R.
TITLE Direct Submission
JOURNAL Submitted (09-AUG-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA, rbuell@tigr.org
COMMENT On Jul 11, 2001 this sequence version replaced gi:1404299.
Address all correspondence to:rice@tigr.org

BAC clone OSJNB0093E13 is from *Oryza sativa* chromosome 3
The orientation of the sequence is from SP6 to T7 end of the BAC
clone.
Genes were identified by a combination of several methods: Gene
prediction programs including fgenesh (<http://www.softberry.com/>),
GENSCAN+ (Chris Burge,
<http://CCR-081.mit.edu/GENSCAN.html>), GeneMarkHMM (Mark Borodovsky,
<http://genemark.biology.gatech.edu/GeneMark/>), and Gensplice
(Mihela Pertea and Steven Salzberg, contact.mperce@tigr.org),
searches of the complete sequence against a peptide database and
the plant EST database at TIGR (<http://www.tigr.org/tdb/tgi.shtml>).
Annotated genes are named to indicate the level of evidence for
their annotation. Genes with similarity to other proteins are named
after the database hits. Genes without significant peptide
similarity but with EST similarity are named as unknown proteins.
Genes without protein or EST similarity, that are predicted by more
than two gene prediction programs over most of their length are
annotated as hypothetical proteins. Genes encoding tRNAs are
predicted by tRNAscan-SE (Sean Eddy,
<http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats are
identified by RepeatMasker (Arian Smit,
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>).
This BAC overlaps with rice BACs OSJNB0091J19 (GB:AC084320) and
OSJNB0024U04.

FEATURES
source Location/Qualifiers
1..129778
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/sub_species="japonica"
/db_xref="taxon:39947"
/chromosome="3"
/map="near S13014"
/clone="OSJNB0093E13"
1701..1821
/rpt_family="Gaijin Oes2 MITE element from gb:D10985 Rice
seirine cbp3 gene for carboxypeptidase III (1125 to 1251)
127 nt
complement(2452..2476)
/rpt_family="(GAA)n"
4620..5435
/gene="OSJNB0093E13.4"
/note="predicted by fgenesh"
join(<4620..4718,4866..>5435)
/gene="OSJNB0093E13.4"
join(4620..4718,4866..5435)
/gene="OSJNB0093E13.4"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAK72893.1"
/db_xref="GI:14718315"
/translation="MEDWVLSGSDSDSVELHDGVSNGSDSGFAFSGKLPSSSQ
LILTLNCKSTQPKRTRRTLPVSPALAGGHEHNAEISDAVERNNHDVDSNIAAEITTF
SGEEDLADTDIDECFDEBDEIDICENPDEIDPDEBDEIDDDSDIDDDICEDYDEDK
ICEENPDDEIDFDEEBIDCEBDDDDDDCESLDDNNICFPADIEDIICLETRTKMSST
NIAM"
4640..4775
/rpt_family="(CGG)n"
complement(4730..4835)
/rpt_family="(CGG)n"
11008..11724
/gene="OSJNB0093E13.5"
/note="predicted by fgenesh"
11008..>11724
/gene="OSJNB0093E13.5"
/rpt_family="(CGG)n"

CDS 11008..11724
/gene="OSJNB0093E13.5"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAK72895.1"
/db_xref="GI:14718317"
/translation="MEVRRRAVAVSESGRRRQAYGEOROSGRROLATGRLLALPEA
KGRWSEGEKATVSGRALGDSGSSASGSGRGRARVGRVPTAPAPVSV
FGDGRROEMGGNGVVBARKTLLIMGRDLSLGIPIGWGRPGGIDIGLRQGW
BAGAVKNAAGAHGRORSSSGGKKEKGRVERRGSTMPFNGRSGREMLMHLVLDASA
QMGQAGHGHLRGACQG"
12209..12383
/rpt_family="(TACA)n"
complement(12766..16846)
/gene="OSJNB0093E13.7"
/note="highly similar to retrofit gag pol protein
gb:AA82754 GI:2586082 (Oryza longistylis)"
complement(join(<12766..12903,13103..13518,13741..>16846))
/gene="OSJNB0093E13.7"
complement(join(12766..12903,13103..13518,13741..16846))
/codon_start=1
/product="putative gag-pol protein"
/protein_id="AAK72882.1"
/db_xref="GI:14718304"
/translation="MASSSSSTAFANPLFGVQVSEKLTQKHSLMSAQVLTALRGARL
EGHVTGTPPEALEEKEGKEKTVAVPNPAVGEWPTADQVGLPSSLTRERS
OVAGPPTAAAKMTIEMPTSRSHAGINVRLLTTQGSVTVEYLSKMALEDGI
NATGRTDDEELVATITNGLDSEFPVYBALAKRASVAVSVSGLGFEKVKIRT
ACATVSGRSGGNGQGRGNPRGRGTGSGGSGRGGHGRNGRGGTNNKPTQVC
HKKGIVVADCMWRVDENYVPDEKLGAAVTHAVGVDTNNVVDTEADH1TGQVLDKTR
EYKGTQD1HTASGEGMDIOHIGHSYVPTSRPLHLNHLVPKASKXLI1SVHRLVAD
NYAFLEIHKQFELIKDQVTRTILEGRCRDLVPLPADPTIKQVAVMSPFRMGRGL
GASKEIVLRVINOMKLECSNSESBSQCDACQCKSHOLPEPKFVSNNLEI1HS
DVMGPASISVCAKRTYVFDIDYKFTWITPLKFSBYPEKFEQSNVBERPKKIL
GMQTDWGEYQLNSFFKQIGIAHQSVPHTHQQAVERKRRH1VEGLSLALASAM
PLKFWDEAFALATYLIINLRPRKVIDFDPLRLFPQDPDYSLRTFGCAWENLRPN
THKQFRSKQCVLFGFSLNHLKFGKLEVAQCQVYSRVTPEOCFPPSKLOPNAGAR
LRAEISLIPADLTSHASDQGVAVSHNDHMHNI1LPVYTSFPENDAADTDASNAEI
NAEOPDGIHPASGTGCVHTAEVVAESSTPDRAPVAPVAPVTPPAATSAVAV
SSRASRS1ASHQIGDTSPPPHHDAAGSSAAPEVIAATEBGDINAPKTRLOSGVR
KKKITYDTGVRKGCFTSSGEFONLAELAGDNKWEKMEKETEYALKMKTWH1VPRKG
INI1DCKNVYKIKRADGSLDRYKGRILAKFKQRYGDIYEBTESPVKAAVIRVLS
LVASKDTS1PFYDKKGVTFM1VVDV1IVSSSSKATSLAD1QDNLQFALDLDLHD
FPLGEYKKNVNDGLVLTQEKYASDPLRVNGSDCKPVNTPISSKLSOGGALGPN
DSTQVRS1VGA1OYVLTLRPD1SPSVNNDP1LYLGCHNKXIQF1DQV1KLNTRHMQO
QOKLCGYKHYFRNLV"
complement(13497..13708)
/rpt_family="Rice retrotransposon Tos7 DNA for reverse
transcriptase, partial cds gi|1621482|dbj|D85871.1|D85871"
complement(16049..16146)
/rpt_family="(CGG)n"
17933..19954
/gene="OSJNB0093E13.6"
/note="predicted by fgenesh"
17933..>19954
/gene="OSJNB0093E13.6"
17933..19954
/gene="OSJNB0093E13.6"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAK72884.1"
/db_xref="GI:14718306"
/translation="MEDWVLSGSDSDSVELHDGVSNGSDSGFAFSGKLPSSSQ
LILTLNCKSTQPKRTRRTLPVSPALAGGHEHNAEISDAVERNNHDVDSNIAAEITTF
SGEEDLADTDIDECFDEBDEIDICENPDEIDPDEBDEIDDDSDIDDDICEDYDEDK
ICEENPDDEIDFDEEBIDCEBDDDDDDCESLDDNNICFPADIEDIICLETRTKMSST
NIAM"
4640..4775
/rpt_family="(CGG)n"
complement(4730..4835)
/rpt_family="(CGG)n"
11008..11724
/gene="OSJNB0093E13.5"
/note="predicted by fgenesh"
11008..>11724
/gene="OSJNB0093E13.5"
/rpt_family="(CGG)n"

	gene	ELTTGTEIPYEVTLPICAAAGVQKGLRPITPKNAHKLSLLCKMCQOQEAPAKPDPFS ILETTLORIAPEVEDHDGHKEKILIGLTSALRGRRH" 65979..84215 /gene="P0705B11.8"					
	mRNA	join(655979..66134,66228..66279,67599..75119,75210..757533,.76241,76395,.76475,76728..77036,78187..78573,78721..79524,79613..79975,80194..80424,80691..81035,81251..81469,81516..81572,81735..82079,82295..82513,					
Query Match	70.7%;	Score 21.2;	DB 8;	Length 140729;			
Best Local Similarity	88.5%;	Pred. No. 2.2e+02;					
Matches 23;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;			
Oy	5	CTTGTCACCGCTTCGTCTGTGCCA 30					
Db	55011	CATGTTAACGCTTCGTCTATTCCA 55036					
RESULT 14							
LOCUS AP005710							
DEFINITION Oryza sativa (japonica cultivar-group) chromosome 9 clone P0646B07,							
ACCSSION AP005710	153675 bp	DNA	linear	HTG 06-SEP-2002			
VERSION AP005710.1	GI:22775481						
KEYWORDS HTG; HTGS PHASE2.							
SOURCE Oryza sativa (japonica cultivar-group)							
ORGANISM Oryza sativa (japonica cultivar-group)							
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoidae; Oryzaceae; Oryza.							
REFERENCE 1							
AUTHORS Sasaki,T., Matsumoto,T. and Katayose,Y.							
TITLE Oryza sativa nipponbare(GAS) genomic DNA, chromosome 9, PAC clone:P0646B07							
JOURNAL Published Only in Database (2002)							
REFERENCE 2 (bases 1 to 153675)							
AUTHORS Sasaki,T., Matsumoto,T. and Katayose,Y.							
TITLE Direct Sublibrary							
JOURNAL Submitted (05-SEP-2002) Takuji Sasaki, National Institute of Agrobiological Sciences, Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan							
(E-mail:tsasakiat@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/, Tel:81-298-38-7441, Fax:81-298-38-7468)							
NOTE: It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have provided by the submitter. This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.							
* NOTE: This is a 'working draft' sequence.							
* This sequence will be replaced							
* by the finished sequence as soon as it is available and							
* the accession number will be preserved.							
FEATURES							
source							
1. 153675							
/organism="Oryza sativa (japonica cultivar-group)"							
/mol_type="genomic DNA"							
/cultivar="Nipponbare"							
/db_xref="taxon:39947"							
/chromosome="9"							
/clone="P0646B07"							
ORIGIN							
Query Match	70.7%;	Score 21.2;	DB 2;	Length 153675;			
Best Local Similarity	88.5%;	Pred. No. 2.2e+02;					
Matches 23;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;			
Oy	5	CTTGTCACCGCTTCGTCTGTGCCA 30					
Db	19380	CATGTTAACGCTTCGTCTATTCCA 19405					

misc_feature 18062..18190,18223..18322,18729..18861,18967..19111))
/gene="OJ1148_D05.5"
complement(join(15074..15176,17827..17877,17931..18029,
18062..18190,18223..18322,18729..18861,18967..19111))
/gene="OJ1148_D05.5"
/note="hypothetical ORF
predicted by GENSCAN
this category is not included in IRGSP standard"
complement(25292..25586)
/gene="OJ1148_D05.6"
complement(join(<25292..25412,25471..>25586))
/gene="OJ1148_D05.6"
/note="start and end point are not identified"
complement(join(25292..25412,25471..25586))
/gene="OJ1148_D05.6"
/note="predicted by GeneMark.hmm etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD07688.1"
/db_xref="GI:41053262"
/translation="MAHGSAGLKAATVTGDMARQRTGREALMRRRRRDCRRRVG
DGAATATRHDAARQWQSEATATAPVPAGLEAG"
complement(join(26131..26319,28092..28139))
/gene="OJ1148_D05.7"
complement(join(26131..26319,28092..28139))
/gene="OJ1148_D05.7"
/note="hypothetical ORF
predicted by GENSCAN
this category is not included in IRGSP standard"
complement(30668..30991)
/gene="OJ1148_D05.8"
complement(30668..30991)
/gene="OJ1148_D05.8"
/note="start and end point are not identified"
complement(30668..30991)
/gene="OJ1148_D05.8"
/note="predicted by FGENSEH etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD07689.1"
/db_xref="GI:41053263"
/translation="MPCTSPSCATTCRTPGNHHGTTSELRAASRRASLWEVLP
VNPPLSPSPYFPSSSCQHLVAMGVAVNTKVVGEGRKIKREEDDEWLNGSHA
ILDPT"
32784..35862
/gene="OJ1148_D05.9"
join(32784..33234,33666..34250,35333..35862)
/gene="OJ1148_D05.9"
/note="supported by full-length cDNA(s) : AK060695"
join(32872..33234,33666..34250,35333..35506)
/gene="OJ1148_D05.9"
/note="contains full-length cDNA(s) : AK060695"
/codon_start=1
/product="putative basic-helix-loop-helix transcription
factor"
/protein_id="BAD07690.1"

Query Match 70.7%; Score 21.2; DB 8; Length 155146;
Best Local Similarity 88.5%; Pred. No. 2.2e+02;
Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 5 CTTGTTCAACCGTTCGTTGTTCA 30
| ||||| ||||| ||||| ||||| |||||
DB 153758 CATGTTAACGTTCTCTATTCCA 153783

Search completed: September 16, 2005, 03:01:38
Job time : 102.667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 08:10:17 ; Search time 14.5242 Seconds

(Without alignment)

Title: US-10-713-381-2_COPY_1179_1208

Perfect score: 30
Sequence: 1 catgctgtctcaacgctcgtctgtccca 30

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: genebegn1980s:*
2: genebegn1980s:*
3: genebegn2000s:*
4: genebegn2001as:*
5: genebegn2001bs:*
6: genebegn2002as:*
7: genebegn2002bs:*
8: genebegn2003as:*
9: genebegn2003bs:*
10: genebegn2003cs:*
11: genebegn2003ds:*
12: genebegn2004as:*
13: genebegn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	100.0	30	AAH76335	Aah76335 Z. mays M
2	30	100.0	255	AAH76340	Aah76340 Z. mays M
3	30	100.0	1394	AAH76348	Aah76348 Zea mays
4	30	100.0	1394	AAH76349	Aah76349 Zea mays
5	30	100.0	1394	AAH76332	Aah76332 Z. mays M
6	30	100.0	1394	AAH76333	Aah76333 Z. mays M
7	24.2	80.7	158	AAH76334	Aah76334 Z. mays M
8	22	73.3	1304	ABO75351	Abq75351 Human lun
9	20.4	68.0	1542	ABZ13819	Abz13819 Arabidops
10	20.2	67.3	2357	ACA44013	Aca44013 Prokaryot
11	20	66.7	2000	ADJ40785	Adj40785 Plant CDN
12	19.6	65.3	2000	ADA72061	Ada72061 Rice gene
13	19.6	65.3	2000	ADA71994	Ada71994 Rice gene
14	19.6	65.3	2000	ADA71736	Ada71736 Rice gene
15	19.6	65.3	2000	ADA72530	Ada72530 Rice gene
16	19.6	65.3	2000	ADJ40977	Adj40977 Plant CDN
17	19.4	64.7	2240	AAQ47668	Aaq47668 Sequence
18	19.4	64.7	2243	AAF29956	Aaf29956 Rat brain
19	19.4	64.7	2243	ACD26214	Acd26214 cDNA enco
20	19.4	64.7	349980	ABQ81845	Abq81845 Bifidobac

C 21	19.2	64.0	2000	8	ADA73340	Ada73340 Rice gene
C 22	19	63.3	1324	8	ADA71777	Ada71777 Rice gene
C 23	19	63.3	1388	3	AAC44356	Aac44356 Arabidops
C 24	19	63.3	1821	10	ADE93879	Ade93879 Ixodes sc
C 25	19	63.3	5671	4	AAK90190	Aak90190 Human dig
C 26	19	63.3	5671	5	ABA20915	Abaz0915 Human ner
C 27	19	63.3	5671	9	AA639832	Aa639832 Genomic B
C 28	19	63.3	5671	9	ADB32792	Adb32792 Human nov
C 29	19	63.3	198522	11	ACN44010	Acn44010 Human gen
C 30	18.8	62.7	883	6	ABK65288	Abk65288 Arabidops
C 31	18.8	62.7	883	10	ADD30938	Add30938 Plant yle
C 32	18.8	62.7	883	10	ADE37244	Ade37244 Plant yle
C 33	18.8	62.7	883	12	ADI41918	Adi41918 Plant tra
C 34	18.8	62.7	1083	4	AA553644	Aa553644 Helicobac
C 35	18.8	62.7	1083	8	ACA34769	Aca34769 Prokaryot
C 36	18.8	62.7	1899	8	ABX06821	Abx06821 S. pneumo
C 37	18.8	62.7	1902	4	AA555583	Aa555583 Streptoco
C 38	18.8	62.7	1902	8	ACA49918	Aca49918 Prokaryot
C 39	18.8	62.7	1908	12	ADJ63835	Adj63835 Plant lyp
C 40	18.8	62.7	2000	12	ADJ41535	Adj41535 Plant CDN
C 41	18.8	62.7	2247	13	ADS57809	Ad57809 Bacterial
C 42	18.8	62.7	2920	11	ADM02356	Adm02356 Human CDN
C 43	18.8	62.7	3339	10	ACF79841	Acf79841 Mouse ani
C 44	18.8	62.7	7812	3	AA61238	Aa61238 DNA enco
C 45	18.8	62.7	9541	2	AAV52265	Aav52265 Streptoco

ALIGNMENTS

RESULT 1
ID AAH76335 standard; DNA; 30 BP.
AC AAH76335;
XX
XX
DT 29-OCT-2001 (first entry)
XX
XX
DE Z. mays Ms45 male tissue-preferred regulatory region fragment.
XX
XX
KM Ms45; male tissue; regulatory region; transcription; male fertility;
KW hybrid seed; de.
XX
XX
OS Zea mays.
XX
XX
PN WO200160997-A2.
XX
XX
PD 23-AUG-2001.
XX
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
XX
PR 15-FEB-2000; 2000US-00504487.
XX
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
XX
DR WPI; 2001-514772/56.
XX
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
XX
XX
PT essential for initiating transcription of the Ms45 gene useful for
XX
XX
PT mediating fertility in a male plant.
XX
XX
PS Claim 5; Page 47; 50pp; English.
XX
XX
CC The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the Ms45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (II) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a DNA fragment -152 to -181 bases upstream of

CC the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
 CC nucleotide sequence
 XX
 SQ Sequence 30 BP; 4 A; 9 C; 5 G; 12 T; 0 U; 0 Other;

Query Match 100.0%; Score 30; DB 5; Length 30;
 Best Local Similarity 100.0%; Pred. No. 0.004;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGCTTGTTCACCGTTCGTTGTTCCA 30
 |||
 DB 1 CATGCTTGTTCACCGTTCGTTGTTCCA 30

RESULT 2

AAH76340
 ID AAH76340 standard; DNA; 255 BP.

AC AAH76340;

DT 29-OCT-2001 (first entry)

DE Z. mays Ms45 promoter fragment.

XX Ms45; male tissue; regulatory region; transcription; male fertility;

KM hybrid seed; promoter; de.

XX Zea mays.

PN WO200160997-A2.

PD 23-AUG-2001.

PF 13-FEB-2001; 2001WO-US004527.

PR 15-FEB-2000; 2000US-00504487.

PA (PION-) PIONEER HI-BRED INT INC.

PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;

DR WPI; 2001-514772/56.

XX A male tissue-preferred regulatory region comprising nucleotide sequences
 PT essential for initiating transcription of the Ms45 gene useful for
 PT mediating fertility in a male plant.

XX Example 5; Fig 8; 50pp; English.

CC The invention provides a male tissue-preferred regulatory region (I)
 CC comprising nucleotide sequences essential for initiating transcription of
 CC the Ms45 gene. A method of mediating male fertility in a plant is
 CC provided that involves introducing an expression vector comprising a
 CC promoter operably linked to (I) into a plant where the exogenous gene
 CC impacts male fertility of the plant and (I) controls expression of the
 CC exogenous gene. A method of producing hybrid seeds is also provided. The
 CC present sequence represents a Z. mays Ms45 promoter fragment

XX Sequence 255 BP; 59 A; 86 C; 39 G; 71 T; 0 U; 0 Other;

Query Match 100.0%; Score 30; DB 5; Length 255;
 Best Local Similarity 100.0%; Pred. No. 0.0051;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGCTTGTTCACCGTTCGTTGTTCCA 30
 |||
 DB 39 CATGCTTGTTCACCGTTCGTTGTTCCA 68

RESULT 3

AAK07408
 ID AAK07408 standard; DNA; 1394 BP.

AC AAK07408;
 XX
 DT 08-JUN-1999 (first entry)

DE Zea mays Ms45 male tissue-preferred regulatory region.

XX Ms45; male; tissue-preferred; regulatory region; plant cells;
 KM plant tissue; differentiated; maize; hybrid seed; fertility; ss.

XX Zea mays.

PN WO9859061-A1.

PD 30-DEC-1998.

PF 19-JUN-1998; 98WO-US012895.

PR 23-JUN-1997; 97US-00880499.

PA (PION-) PIONEER HI-BRED INT INC.

PI Albertsen MC, Fox TW, Garnaat CW, Huffman GA, Kendall TL;

DR WPI; 1999-105628/09.

PT New nucleic acid encoding a Ms45 male tissue-preferred regulatory region
 PT - useful in mediating plant fertility, especially hybrid seed production.

XX Claim 2; Page 22-23; 39pp; English.

XX The sequence is that encoding an Ms45 male tissue-preferred regulatory
 CC region. It may be used in the construction of a vector for a method of
 CC producing exogenous genes in a male tissue- preferred manner, which is
 CC useful in restoring or conferring fertility, such as in hybrid seed
 CC production. In conferring fertility, a monocot/dicot plant is transformed
 CC with the exogenous nucleotide sequence (a male sterility gene, preferably
 CC Ms45), which encodes a product selected from auxins, rolB and diptheria
 CC toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
 CC and infertile plants

XX SQ Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;

Query Match 100.0%; Score 30; DB 2; Length 1394;
 Best Local Similarity 100.0%; Pred. No. 0.0063;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGCTTGTTCACCGTTCGTTGTTCCA 30
 |||
 DB 1179 CATGCTTGTTCACCGTTCGTTGTTCCA 1208

RESULT 4

AAK07409
 ID AAK07409 standard; DNA; 1394 BP.

AC AAK07409;

DT 08-JUN-1999 (first entry)

DE Zea mays Ms45 male tissue-preferred regulatory region.

XX Ms45; male; tissue-preferred; regulatory region; plant cells;
 KM plant tissue; differentiated; hybrid seed; fertility; ss.

XX Zea mays.

PN WO9859061-A1.

PD 30-DEC-1998.

PF 19-JUN-1998; 98WO-US012895.

PR 23-JUN-1997; 97US-00880499.

```
XX (PION-) PIONEER HI-BRED INT INC.
PA
XX
XX Albertsen MC, Fox TW, Garnaat CW, Huffman GA, Kendall TL;
XX
XX WPI; 1999-105628/09.
XX
XX New nucleic acid encoding a Ms45 male tissue-preferred regulatory region
PT - useful in mediating plant fertility, especially hybrid seed production.
XX
XX Claim 3; Page 23-24; 39pp; English.
XX
XX The sequence is that encoding an Ms45 male tissue-preferred regulatory
CC region. It may be used in the construction of a vector for a method of
CC producing exogenous genes in a male tissue-preferred manner, which is
CC useful in restoring or conferring fertility, such as in hybrid seed
CC production. In conferring fertility, a monocot/dicot plant is transformed
CC with the exogenous nucleotide sequence (a male sterility gene, preferably
CC Ms45), which encodes a product selected from auxins, rolB and diptheria
CC toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
CC and infertile plants
XX
XX Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;
SQ
XX
XX Query Match 100.0%; Score 30; DB 2; Length 1394;
XX Best Local Similarity 100.0%; Pred. No. 0.0063;
XX Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CATGCTTGTTCACCGTTCGTTGTTCCA 30
DB 1179 CATGCTTGTTCACCGTTCGTTGTTCCA 1208
XX
XX RESULT 5
XX AAH76332
XX ID AAH76332 standard; DNA; 1394 BP.
XX
XX AAH76332;
XX
XX 29-OCT-2001 (first entry)
XX
XX Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.
DE
XX Ms45; male tissue; regulatory region; transcription; male fertility;
XX hybrid seed; ds.
XX
XX Zea mays.
OS
XX WO200160997-A2.
XX
XX 23-AUG-2001.
XX
XX 13-FEB-2001; 2001WO-US004527.
XX
XX 15-FEB-2000; 2000US-00504487.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
XX WPI; 2001-514772/56.
XX
XX A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the Ms45 gene useful for
PT mediating fertility in a male plant.
XX
XX Claim 4; Page 46; 50pp; English.
XX
XX The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the Ms45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a nucleic acid sequence encoding an Ms45 male
CC tissue preferred regulatory region from Z. mays
XX
XX Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;
SQ
XX
XX Query Match 100.0%; Score 30; DB 5; Length 1394;
XX Best Local Similarity 100.0%; Pred. No. 0.0063;
XX Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CATGCTTGTTCACCGTTCGTTGTTCCA 30
DB 1179 CATGCTTGTTCACCGTTCGTTGTTCCA 1208
XX
XX RESULT 7
```

```
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a nucleic acid sequence encoding an Ms45 male
CC tissue preferred regulatory region from Z. mays
XX
XX Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;
SQ
XX
XX Query Match 100.0%; Score 30; DB 5; Length 1394;
XX Best Local Similarity 100.0%; Pred. No. 0.0063;
XX Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CATGCTTGTTCACCGTTCGTTGTTCCA 30
DB 1179 CATGCTTGTTCACCGTTCGTTGTTCCA 1208
XX
XX RESULT 6
XX AAH76333
XX ID AAH76333 standard; DNA; 1394 BP.
XX
XX AAH76333;
XX
XX 29-OCT-2001 (first entry)
XX
XX Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.
DE
XX Ms45; male tissue; regulatory region; transcription; male fertility;
XX hybrid seed; ds.
XX
XX Zea mays.
OS
XX WO200160997-A2.
XX
XX 23-AUG-2001.
XX
XX 13-FEB-2001; 2001WO-US004527.
XX
XX 15-FEB-2000; 2000US-00504487.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
XX WPI; 2001-514772/56.
XX
XX A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the Ms45 gene useful for
PT mediating fertility in a male plant.
XX
XX Claim 4; Page 47; 50pp; English.
XX
XX The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the Ms45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a nucleic acid sequence encoding an Ms45 male
CC tissue preferred regulatory region from Z. mays
XX
XX Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;
SQ
XX
XX Query Match 100.0%; Score 30; DB 5; Length 1394;
XX Best Local Similarity 100.0%; Pred. No. 0.0063;
XX Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CATGCTTGTTCACCGTTCGTTGTTCCA 30
DB 1179 CATGCTTGTTCACCGTTCGTTGTTCCA 1208
XX
XX RESULT 7
```

```

AAH76334
ID   AAH76334 standard; DNA; 158 BP.
XX
XX   AAH76334;
AC
XX
XX   29-OCT-2001 (first entry)
XX
XX   Z. mays Ms45 male tissue-preferred regulatory region fragment.
DE
XX
XX   Ms45; male tissue; regulatory region; transcription; male fertility;
KM   hybrid seed; db.
XX
XX   Zea mays.
OS
XX
XX   MO200160997-A2.
PN
XX
XX   23-AUG-2001.
PD
XX
XX   13-FEB-2001; 2001WO-US004527.
XX
XX   15-FEB-2000; 2000US-00504487.
XX
XX   (PION-) PIONEER HI-BRED INT INC.
PA
XX
XX   Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
PI
XX
XX   WPI; 2001-514772/56.
DR
XX
XX   A male tissue-preferred regulatory region comprising nucleotide sequences
PT   essential for initiating transcription of the Ms45 gene useful for
PT   mediating fertility in a male plant.
XX
XX   Claim 5; Page 47; 50pp; English.
PS
XX
XX   The invention provides a male tissue-preferred regulatory region (I)
CC   comprising nucleotide sequences essential for initiating transcription of
CC   the Ms45 gene. A method of mediating male fertility in a plant is
CC   provided that involves introducing an expression vector comprising a
CC   promoter operably linked to (i) into a plant where the exogenous gene
CC   impacts male fertility of the plant and (ii) controls expression of the
CC   exogenous gene. A method of producing hybrid seeds is also provided. The
CC   present sequence represents a DNA fragment -38 to -195 bases upstream of
CC   the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
CC   nucleotide sequence
XX
XX   Sequence 158 BP; 41 A; 50 C; 21 G; 46 T; 0 U; 0 Other;
SQ
XX
XX   Query Match      80.7%; Score 24.2; DB 5; Length 158;
XX   Best Local Similarity 89.7%; Pred. No. 1.5;
XX   Matches 26; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY
XX
XX   1 CATGCTGTTCACCGTGTGCTGTTC 29
DB
XX
XX   25 CATGCTGTTCACCGTGTGCTGTTC 53

```

```

XX
XX   22-AUG-2002.
PD
XX
XX   20-NOV-2001; 2001WO-US045080.
XX
XX   PF
XX
XX   20-NOV-2000; 2000US-0252054P.
XX
XX   PR
XX
XX   (DIAD-) DIADEXUS INC.
PA
XX
XX   Macina RA, Recipon H, Chen S, Sun Y, Liu C;
PI
XX
XX   WPI; 2002-657601/70.
DR
XX
XX   New lung specific nucleic acid useful in gene therapy or as vaccines for
PT   treating lung cancer (e.g. squamous cell carcinoma) or non-cancerous lung
PT   diseases, as well as for diagnosing, monitoring or staging these
PT   diseases.
XX
XX   Claim 1; Page 206; 282pp; English.
XX
XX   The present invention describes an isolated lung specific nucleic acid
CC   (LSNA) comprising a sequence that: (a) encodes any of the 93 amino acid
CC   sequences comprising 17 - 733 amino acids, given in ABP52873 to ABP52965;
CC   (b) comprises any of 115 sequences comprising 148 - 3193 base pairs (bp);
CC   given in ABQ75262 to ABQ75376; (c) selectively hybridizes to (a) or (b);
CC   or (d) has 60 % sequence identity to (a) or (b). LSNA and lung specific
CC   protein (LSP) sequences have cytostatic activity and can be used in gene
CC   therapy and vaccines. LSNA and LSPs are useful for diagnosing and
CC   monitoring the presence and metastases of lung cancer in a patient. An
CC   antibody that specifically binds to an LSP can be used for determining
CC   the presence of an LSP in a sample, as well as for treating a patient
CC   with lung cancer, particularly by inducing an immune response against the
CC   lung cancer cell expressing the LSNA or LSPs. In particular, these LSNA
CC   and LSPs are useful for identifying, diagnosing, monitoring, staging,
CC   imaging and treating lung cancer (e.g. squamous cell carcinoma) and non-
CC   cancerous disease states in lung
XX
XX   Sequence 1304 BP; 352 A; 347 C; 314 G; 291 T; 0 U; 0 Other;
SQ
XX
XX   Query Match      73.3%; Score 22; DB 6; Length 1304;
XX   Best Local Similarity 83.3%; Pred. No. 16;
XX   Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
OY
XX
XX   1 CATGCTGTTCACCGTGTGCTGTTC 30
DB
XX
XX   1155 CTTTCTGTTCGCCCGTTCCTTGTTC 1126

```

```

RESULT 8
ABQ75351/C
ID   ABQ75351 standard; cDNA; 1304 BP.
XX
XX   ABQ75351;
AC
XX
XX   05-NOV-2002 (first entry)
XX
XX   Human lung specific nucleic acid sequence SEQ ID NO:90.
DE
XX
XX   Human; lung; lung specific nucleic acid; LSNA; lung specific protein;
KM   LSP; cytostatic; gene therapy; vaccine; metastasis; lung cancer;
XX   squamous cell carcinoma; gene; ss.
OS
XX
XX   Homo sapiens.
XX
XX   WO200264788-A2.

```

```

RESULT 9
ABZ13819
ID   ABZ13819 standard; DNA; 1542 BP.
XX
XX   ABZ13819;
AC
XX
XX   21-JAN-2003 (first entry)
XX
XX   Arabidopsis thaliana scree regulated gene SEQ ID NO 1624.
DE
XX
XX   Arabidopsis thaliana; plant; gene; stress; transgenic; db.
XX
XX   Arabidopsis thaliana.
OS
XX
XX   WO200216655-A2.
PN
XX
XX   28-FEB-2002.
PD
XX
XX   24-AUG-2001; 2001WO-US026685.
XX
XX   PF
XX
XX   24-AUG-2000; 2000US-0227866P.
XX
XX   PR
XX
XX   26-JAN-2001; 2001US-0264647P.
XX
XX   PR
XX
XX   22-JUN-2001; 2001US-0300111P.
XX
XX   (SCRI ) SCRIPPS RES INST.

```

PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX
 PI Harper JF, Kreps J, Wang X, Zhu T;
 XX
 DR WPI; 2002-304127/34.
 XX
 PT Identifying a stress condition to which a plant cell has been exposed and
 PT producing plants with increased tolerance to these abiotic stresses.
 XX
 PS Claim 144; SEQ ID NO 1624; 577bp + Sequence Listing; English.
 XX
 CC The invention relates to identifying a stress condition to which a plant
 CC cell has been exposed, comprising: (a) contacting nucleic acid
 CC representative of expressed polynucleotides in the plant cell with an
 CC array or probes representative of the plant cell genome; and (b)
 CC detecting a profile of expressed polynucleotides in the plant cell
 CC characteristic of a stress response. The method is useful in the
 CC production of transgenic plants, cells and seeds and in producing plants
 CC with increased tolerance to abiotic stress. The present sequence is that
 CC of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used
 CC in methods of the invention. Note: The sequence data for this patent is
 CC not represented in the printed specification but is based on sequence
 CC information supplied to Derwent by the European Patent Office
 SQ Sequence 1542 BP; 345 A; 382 C; 370 G; 445 T; 0 U; 0 Other;
 Query Match 68.0%; Score 20.4; DB 6; Length 1542;
 Best Local Similarity 80.0%; Pred. No. 81;
 Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 CATGCTGTTCACCGTTCGTTGTTCA 30
 Db 237 CATGATTGTTCACCAATCGTCGTTTCCA 266
 RESULT 10
 ID ACA44013 standard; DNA; 2397 BP.
 AC ACA44013;
 XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Prokaryotic essential gene #25670.
 XX
 KW Antisense; ds; prokaryotic essential gene; cell proliferation;
 KW drug design; gene.
 OS Pseudomonas putida.
 XX
 PN WO20027183-A2.
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362659P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zykelding JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI; 2003-029926/02.
 XX
 DR P-PsDB; AB040143.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.

XX
 PS Claim 14; SEQ ID NO 31883; 1766bp; English.
 CC
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-regulated gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 SQ Sequence 2397 BP; 415 A; 825 C; 781 G; 376 T; 0 U; 0 Other;
 Query Match 67.3%; Score 20.2; DB 8; Length 2397;
 Best Local Similarity 88.0%; Pred. No. 1e+02;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 CATGCTGTTCACCGTTCGTTG 25
 Db 516 CATGCTGTTCACCGTTCGTTG 540
 RESULT 11
 ID ADJ40785 standard; cDNA; 2000 BP.
 AC ADJ40785;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE Plant cDNA #1785.
 XX
 KW Plant; gene; ss; transcription; plant genome augmentation; cereal;
 KW soybean; alfalfa; sunflower; canola; cotton; peanut; tobacco; sugar beet;
 KW maize; barley; sorghum; rice; wheat; crop plant; insecticide resistance;
 KW stress tolerance; salt tolerance; cold tolerance; drought tolerance;
 KW plant nutrition; apical dominance; dwarfism; early flowering; antiviral;
 KW antifungal.
 XX
 OS Eukaryota.
 XX
 PN US2004016025-A1.
 XX
 PD 22-JAN-2004.
 XX
 PF 26-SEP-2002; 2002US-00260238.
 XX
 PR 26-SEP-2001; 2001US-0325277P.
 PR 26-SEP-2001; 2001US-0325448P.

```

PR 04-APR-2002; 2002US-0370620P.
XX
XX (BUDW/) BUDWORTH P.
PA (MOUG/) MOUGHAMER T.
PA (BRIG/) BRIGGS S P.
PA (COOP/) COOPER B.
PA (GLAZ/) GLAZEBROOK J.
PA (GOFF/) GOFF S A.
PA (KATA/) KATAGIRI F.
PA (KREP/) KREPS J.
PA (PROV/) PROVART N.
PA (RICK/) RICE D.
PA (ZHUT/) ZHU T.
XX
XX Budworth P, Moughamer T, Briggs SP, Cooper B, Glazebrook J;
PI Goff SA, Katagiri F, Krepes J, Provart N, Riche D, Zhu T;
XX
XX WPI; 2004-190374/18.
XX
XX New rice promoter, useful for manipulating crop plants to alter or
PT improve phenotypic characteristics, e.g. produce large quantities of oil
PT or proteins, resistance to insecticides, virus or fungi, stress tolerance
PT or high nutritional value.
XX
XX Claim 26; SEQ ID NO 1785; 230pp; English.
XX
XX The invention relates to plant nucleotide sequences that direct seed-,
CC leaf- and/or stem-, panicle-, root- or pollen-specific or -preferential
CC or constitutive transcription of an operatively linked nucleic acid
CC segment. The invention also relates to a method for augmenting a plant
CC genome and a method of identifying a gene, where its expression is
CC altered in the seed, leaf, stem, panicle, pollen, root or is constitutive
CC in a plant cell. The plant is a cereal, e.g. soybean, alfalfa, sunflower,
CC canola, cotton, peanut, tobacco or sugar beet, preferably maize, barley,
CC sorghum, rice or wheat. The polynucleotides and the polypeptides they
CC encode are useful for manipulating crop plants to alter or improve
CC phenotypic characteristics, to produce large quantities of oil or
CC proteins, to incur resistance to insecticides, viruses or fungi, and to
CC incur stress tolerance (e.g. salt, cold or drought) to ensure the plants
CC have a high nutritional value with reduced apical dominance or dwarfism,
CC early flowering or altered metabolic pathways. This sequence represents a
CC plant nucleic acid of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification but was obtained in
CC electronic format directly from USPTO at segdata.uspto.gov/sequence.html.
XX
XX Sequence 2000 BP; 552 A; 438 C; 444 G; 566 T; 0 U; 0 Other;
SQ
XX
XX Query Match 66.7%; Score 20; DB 12; Length 2000;
XX Best Local Similarity 82.1%; Pred. No. 1.2e+02;
XX Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
XX
XX 3 TGCTGTCAACCGTCGTCTGTCCA 30
DB 1179 TACATGTTAAACCGTCGTCTTATTCAA 1206
XX
XX RESULT 12
XX ADA72061
XX ID ADA72061 standard; DNA; 2000 BP.
XX
XX ADA72061;
XX
XX 20-NOV-2003 (first entry)
XX
XX Rice gene, SEQ ID 5386.
XX
XX Plant; bacterial infection; fungal infection; viral infection; rice;
XX gene; ds.
XX
XX Oryza sativa.
XX
XX WO2003000898-A1.
XX
XX PN
XX PT

```

```

PD 03-JAN-2003.
XX
XX 22-JUN-2001; 2001WO-1B001105.
XX
XX 22-JUN-2001; 2001WO-1B001105.
XX
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX
XX WPI; 2003-175290/17.
XX
XX Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
XX
XX Claim 27; SEQ ID NO 5386; 899pp; English.
XX
XX The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC illustrate the invention.
XX
XX Sequence 2000 BP; 605 A; 377 C; 326 G; 690 T; 0 U; 2 Other;
SQ
XX
XX Query Match 65.3%; Score 19.6; DB 8; Length 2000;
XX Best Local Similarity 84.6%; Pred. No. 1.8e+02;
XX Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
XX
XX 5 CTGTGTCAACCGTCGTCTGTCCA 30
DB 974 CATGTTAAACCGTCGTCTTATTCAA 999
XX
XX RESULT 13
XX ADA71994/C
XX ID ADA71994 standard; DNA; 2000 BP.
XX
XX ADA71994;
XX
XX 20-NOV-2003 (first entry)
XX
XX Rice gene, SEQ ID 5319.
XX
XX Plant; bacterial infection; fungal infection; viral infection; rice;
XX gene; ds.
XX
XX Oryza sativa.
XX
XX WO2003000898-A1.
XX
XX 03-JAN-2003.
XX
XX 22-JUN-2001; 2001WO-1B001105.
XX
XX 22-JUN-2001; 2001WO-1B001105.
XX
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX
XX WPI; 2003-175290/17.
XX
XX Identifying at least one gene involved in plant resistance or response to
PT

```


PT	pathogenic infection for conferring resistance or tolerance to a plant to
PT	bacterial, fungal or viral infection by determining or detecting plant
PS	gene expression.
XX	
XX	Claim 27; SEQ ID NO 5319; 899bp; English.
CC	The present invention relates to a method (M1) for identifying genes
CC	involved in plant resistance or response to pathogenic infection. M1
CC	comprises identifying a gene whose expression is significantly altered in
CC	the incompatible interaction of plant gene expression relative to
CC	expression of the gene in an uninfected plant, in a mutant plant that
CC	does not express a gene associated with response to pathogenic infection,
CC	or in a corresponding incompatible or compatible interaction. (M1) is
CC	useful for conferring resistance to resistance or tolerance to a plant to
CC	bacterial, fungal or viral infection. The present sequence was used to
XX	illustrate the invention.
SQ	Sequence 2000 BP; 638 A; 400 C; 347 G; 615 T; 0 U; 0 Other;
Query Match	65.3%; Score 19.6; DB 8; Length 2000;
Best Local Similarity	84.6%; Pred. No. 1.8e+02;
Matches	22; Conservative 0; Mismatches 4; Indels 0; Gaps 0
Oy	5 CTGTTCACCGTTCGCTTTGCCA 30 Db 150 CATGTTAACCGTCGCTTAATCCA 125
RESULT 14	
ADA71736/c	
ID ADA71736 standard; DNA; 2000 BP.	
XX	
AC ADA71736;	
XX	
DT 20-NOV-2003 (first entry)	
DE Rice gene, SEQ ID 5061.	
XX	
KW Plant; bacterial infection; fungal infection; viral infection; rice;	
XX	gene; ds.
OS Oryza sativa.	
PN WO200300898-A1.	
XX	
PB 03-JAN-2003.	
XX	
FP 22-JUN-2001; 2001WO-IB001105.	
XX	
PR 22-JUN-2001; 2001WO-IB001105.	
XX	
PA (SYGN) SYNGENTA PARTICIPATIONS AG.	
XX	
E1 Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;	
F1 Katsagiri F, Quan S, Tao Y, Whitcham S, Xie Z, Zhu T, Zou G;	
XX	
DR WPI; 2003-175290/17.	
PT Identifying at least one gene involved in plant resistance or response to	
PT pathogenic infection for conferring resistance or tolerance to a plant to	
PT bacterial, fungal or viral infection by determining or detecting plant	
PT gene expression.	
PS Claim 27; SEQ ID NO 5061; 899bp; English.	
XX	
CC The present invention relates to a method (M1) for identifying genes	
CC involved in plant resistance or response to pathogenic infection. M1	
CC comprises identifying a gene whose expression is significantly altered in	
CC the incompatible interaction of plant gene expression relative to	
CC expression of the gene in an uninfected plant, in a mutant plant that	
CC does not express a gene associated with response to pathogenic infection,	
CC or in a corresponding incompatible or compatible interaction. (M1) is	
CC useful for conferring resistance to resistance or tolerance to a plant to	
CC bacterial, fungal or viral infection. The present sequence was used to	
XX	

```
CC bacterial, fungal or viral infection. The present sequence was used to
XX illustrate the invention.
CC
SQ Sequence 2000 BP; 601 A; 422 C; 306 G; 671 T; 0 U; 0 Other;

Query Match          65.3%; Score 19.6; DB 8; Length 2000;
Best Local Similarity 84.6%; Pred. No. 1.8e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      5 CTGTTCACCGTTCGCTTTGCCA 30
        |||||
DB       291 CATGTTAACCGTGCTTAATCA 266

RESULT 15
ADA72530/C
ID ADA72530 standard; DNA; 2000 BP.
XX
AC ADA72530;
XX
DT 20-NOV-2003 (first entry)
XX
DE Rice gene, SEQ ID 5855.
XX
KW Plant; bacterial infection; fungal infection; viral infection; rice;
KM gene; ds.
XX
OS Oryza sativa.
PN WO200300898-A1.
PP 03-JAN-2003.
PD 22-JUN-2001; 2001WO-IB001105.
PE 22-JUN-2001; 2001WO-IB001105.
PR 22-JUN-2001; 2001WO-IB001105.
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.
PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
DR WPI, 2003-175290/17.
XX
PT Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
PS Claim 27; SEQ ID NO 5855; 899pp; English.
XX
CC The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or whose expression is significantly altered in
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.
XX
SQ Sequence 2000 BP; 549 A; 413 C; 443 G; 588 T; 0 U; 7 Other;

Query Match          65.3%; Score 19.6; DB 8; Length 2000;
Best Local Similarity 84.6%; Pred. No. 1.8e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      5 CTGTTCACCGTTCGCTTTGCCA 30
        |||||
DB       27 CATGTTAACCGTGCTTAATCA 2
```

Search completed: September 15, 2005, 21:33:22
Job time : 15.5242 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using bw model

Run on: September 14, 2005, 19:14:19 ; Search time 4.53786 Seconds
(without alignments)
10817.505 Million cell updates/sec

Title: US-10-713-381-2_COPY_1179_1208

Perfect score: 30
Sequence: 1 catgctgtcaccgtcgtcgtcgtccca 30

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/1/ina/5C_COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/prodata/1/ina/backfillseq1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	100.0	1394	3 US-08-880-499-1	Sequence 1, Appl
2	30	100.0	1394	3 US-08-880-499-2	Sequence 2, Appl
3	19.4	64.7	2243	1 US-07-837-609-15	Sequence 15, Appl
4	19.4	64.7	2243	3 US-08-029-170-15	Sequence 15, Appl
5	19.4	64.7	2243	4 US-09-443-745-15	Sequence 15, Appl
6	19.4	64.7	2243	4 US-09-443-745-15	Sequence 15, Appl
7	18.8	62.7	8756	3 US-09-349-016-1438	Sequence 1, Appl
8	18.8	62.7	8756	4 US-09-349-016-1438	Sequence 1, Appl
9	18.8	62.7	8756	3 US-08-961-527-132	Sequence 132, App
10	18.8	62.7	8756	4 US-09-349-016-13180	Sequence 13180, A
11	18.6	62.0	549	4 US-09-270-767-29480	Sequence 29480, A
12	18.6	62.0	1120	4 US-09-270-767-13492	Sequence 13492, A
13	18.6	61.3	1260	4 US-09-248-796A-3293	Sequence 3293, A
14	18.2	60.7	2710	4 US-09-543-080A-7	Sequence 7, Appl
15	18.2	60.0	601	4 US-09-949-016-200252	Sequence 200252,
16	18.2	60.0	601	4 US-09-949-016-200253	Sequence 200253,
17	18.2	60.0	889	4 US-08-956-171E-1190	Sequence 1190, Ap
18	18.2	60.0	889	4 US-08-956-171E-1190	Sequence 1190, Ap
19	18.2	60.0	11591	4 US-08-728-956-3	Sequence 3, Appl
20	18.2	60.0	11591	4 US-08-728-956-3	Sequence 3, Appl
21	18.2	60.0	254366	4 US-09-949-016-17379	Sequence 17379, A
22	17.8	59.3	333	4 US-09-823-871-3	Sequence 3, Appl
23	17.8	59.3	333	4 US-09-583-110-1388	Sequence 1388, Ap
24	17.8	59.3	333	4 US-09-107-433-743	Sequence 743, Ap
25	17.8	59.3	499	4 US-09-270-767-17275	Sequence 17275, A
26	17.8	59.3	601	4 US-09-270-767-17275	Sequence 17275, A
27	17.8	59.3	601	4 US-09-949-016-142937	Sequence 142937,
				4 US-09-949-016-142937	Sequence 142937,
				4 US-09-949-016-204029	Sequence 204029,

c 28	17.8	59.3	601	4 US-09-949-016-204030	Sequence 204030,
c 29	17.8	59.3	978	3 US-08-858-207A-174	Sequence 174, App
c 30	17.8	59.3	999	3 US-08-858-207A-174	Sequence 174, App
c 31	17.8	59.3	2009	1 US-07-958-222A-1	Sequence 1, Appl
c 32	17.8	59.3	8073	4 US-09-949-016-12327	Sequence 12327, A
c 33	17.8	59.3	8074	4 US-09-949-016-12327	Sequence 12327, A
c 34	17.8	59.3	9472	1 US-08-325-547-9	Sequence 9, Appl
c 35	17.8	59.3	40505	4 US-09-949-016-13439	Sequence 13439, A
c 36	17.8	59.3	72843	4 US-09-949-016-12574	Sequence 12574, A
c 37	17.8	59.3	14115	4 US-09-949-016-17490	Sequence 17490, A
c 38	17.8	59.3	183202	4 US-09-949-016-13614	Sequence 13614, A
c 39	17.8	59.3	265038	4 US-09-949-016-15779	Sequence 15779, A
c 40	17.8	59.3	276237	4 US-09-949-016-17504	Sequence 17504, A
c 41	17.4	58.0	331	4 US-09-640-211A-169	Sequence 169, App
c 42	17.4	58.0	601	4 US-09-949-016-178197	Sequence 178197,
c 43	17.4	58.0	719	4 US-09-107-433-820	Sequence 820, App
c 44	17.4	58.0	789	4 US-09-252-991A-11034	Sequence 11034, A
c 45	17.4	58.0	813	4 US-09-252-991A-5455	Sequence 5455, Ap

ALIGNMENTS

RESULT 1
US-08-880-499-1
Sequence 1, Application US/08880499

Patent No. 6037523

GENERAL INFORMATION:

APPLICANT: Albertson, Marc C.

APPLICANT: Fox, Tim W.

APPLICANT: Carl, Garnat W.

APPLICANT: Kendall, Timmy L.

TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSER: PIONEER HI-BRED INTERNATIONAL, INC.

STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.

STREET: Box 1000

CITY: Johnston

STATE: Iowa

COUNTRY: USA

ZIP: 50131

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/880,499

FILING DATE: CONCURRENTLY HEREWITH

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Sweeney, Patricia A.

REGISTRATION NUMBER: 32,733

REFERENCE/DOCKET NUMBER: 0578

TELECOMMUNICATION INFORMATION:

TELEPHONE: (515) 248-4800

TELEFAX: (515) 248-4844

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1394 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-880-499-1

Query Match 100.0%, Score 30; DB 3; Length 1394;
Best local similarity 100.0%, Pred. No. 0.0009;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/029,170
FILING DATE: 19930310
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/937,609
FILING DATE: 02-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/928,033
FILING DATE: 11-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/861,769
FILING DATE: 01-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/831,248
FILING DATE: 07-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/166 NIHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 2243 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
CLONE: rat brain CCKB receptor
FEATURE:
NAME/KEY: CDS
LOCATION: 136..1494
US-08-029-170-15

Query Match      64.7% Score 19.4; DB 3; Length 2243;
Best Local Similarity 79.3%; Pred. No. 49;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy      1 CATGCTGTTCAACCGTTCGTTGTTCC 29
Db      5 CCTGCTGCTCAACTCTACGCTTGTTC 33

RESULT 5
US-09-443-745-15
Sequence 15, Application US/09443745
Patent No. 6706493
GENERAL INFORMATION:
APPLICANT: WANK, Stephen A.
TITLE OF INVENTION: CLONING AND FUNCTIONAL EXPRESSION OF
TITLE OF INVENTION: CHOLECYSTOKININ RECEPTOR-ENCODING DNA
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSES: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/443,745
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
```

```

APPLICATION NUMBER: 08/029,170
FILING DATE: 10-MAR-1993
APPLICATION NUMBER: US 07/937,609
FILING DATE: 02-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/928,033
FILING DATE: 11-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/861,769
FILING DATE: 01-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/831,248
FILING DATE: 07-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/166 NIHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 2243 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
CLONE: rat brain CCKB receptor
FEATURE:
NAME/KEY: CDS
LOCATION: 136..1494
US-09-443-745-15
```

```

Query Match      64.7% Score 19.4; DB 4; Length 2243;
Best Local Similarity 79.3%; Pred. No. 49;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy      1 CATGCTGTTCAACCGTTCGTTGTTCC 29
Db      5 CCTGCTGCTCAACTCTACGCTTGTTC 33
```

```

RESULT 6
US-09-949-016-16262
Sequence 16262, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CLO01307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 16262
LENGTH: 45571
TYPE: DNA
ORGANISM: Human
US-09-949-016-16262

Query Match      63.3% Score 19; DB 4; Length 45571;
Best Local Similarity 81.5%; Pred. No. 1.3e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 FILE REFERENCE: CLO01307
 CURRENT APPLICATION NUMBER: US/09/949,016
 PRIOR FILING DATE: 2000-04-14
 PRIOR APPLICATION NUMBER: 60/241,755
 PRIOR FILING DATE: 2000-10-20
 PRIOR APPLICATION NUMBER: 60/237,768
 PRIOR FILING DATE: 2000-10-03
 PRIOR APPLICATION NUMBER: 60/231,498
 PRIOR FILING DATE: 2000-09-08

```
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: PairedSeq for Windows Version 4.0
/ SEQ ID NO 13180
/ LENGTH: 112874
/ TYPE: DNA
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)..(112874)
/ OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13180

Query Match          62.7%; Score 18.8; DB 4; Length 112874;
Best Local Similarity 76.7%; Pred. No. 1.9e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 TCCTGTTTCACCGTTCGTTGTTCCA 30
Db 41167 CATGCTTTCGACACAGTCTGCTGCCA 41138

RESULT 11
US-09-270-767-29480/C
/ Sequence 29480, Application US/09270767
/ Patent No. 6703491
/ GENERAL INFORMATION:
/ APPLICANT: Homburger et al.
/ TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
/ FILE REFERENCE: File Reference: 7326-094
/ CURRENT APPLICATION NUMBER: US/09/270,767
/ CURRENT FILING DATE: 1999-03-17
/ NUMBER OF SEQ ID NOS: 62517
/ SOFTWARE: Patentin Ver. 2.0
/ SEQ ID NO 29480
/ LENGTH: 549
/ TYPE: DNA
/ ORGANISM: Drosophila melanogaster
US-09-270-767-29480

Query Match          62.0%; Score 18.6; DB 4; Length 549;
Best Local Similarity 84.0%; Pred. No. 84;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 TGCTTGTTCACCGTTCGTTGTTT 27
Db 223 TGCTTGTTCAGAGTTCCTTCTTGT 199

RESULT 12
US-09-270-767-13492/C
/ Sequence 13492, Application US/09270767
/ Patent No. 6703491
/ GENERAL INFORMATION:
/ APPLICANT: Homburger et al.
/ TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
/ FILE REFERENCE: File Reference: 7326-094
/ CURRENT APPLICATION NUMBER: US/09/270,767
/ CURRENT FILING DATE: 1999-03-17
/ NUMBER OF SEQ ID NOS: 62517
/ SOFTWARE: Patentin Ver. 2.0
/ SEQ ID NO 13492
/ LENGTH: 1120
/ TYPE: DNA
/ ORGANISM: Drosophila melanogaster
US-09-270-767-13492

Query Match          62.0%; Score 18.6; DB 4; Length 1120;
Best Local Similarity 84.0%; Pred. No. 96;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 TGCTTGTTCACCGTTCGTTGTTT 27
Db 223 TGCTTGTTCAGAGTTCCTTCTTGT 199
```

```
RESULT 13
US-09-248-796A-3293
/ Sequence 3293, Application US/09248796A
/ Patent No. 6747137
/ GENERAL INFORMATION:
/ APPLICANT: Keith Weinstein et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
/ FILE REFERENCE: FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 107196.132
/ CURRENT APPLICATION NUMBER: US/09/248,796A
/ CURRENT FILING DATE: 1999-02-12
/ PRIOR APPLICATION NUMBER: US 60/074,725
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR APPLICATION NUMBER: US 60/096,409
/ PRIOR FILING DATE: 1998-08-13
/ NUMBER OF SEQ ID NOS: 28208
/ SEQ ID NO 3293
/ LENGTH: 1260
/ TYPE: DNA
/ ORGANISM: Candida albicans
US-09-248-796A-3293

Query Match          61.3%; Score 18.4; DB 4; Length 1260;
Best Local Similarity 78.6%; Pred. No. 1.2e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 3 TGCTTGTTCACCGTTCGTTGTTCCA 30
Db 903 TGCTTGTTCACACTGTGTTGCCGCCA 930

RESULT 14
US-09-573-080A-7/C
/ Sequence 7, Application US/09573080A
/ Patent No. 6828097
/ GENERAL INFORMATION:
/ APPLICANT: JOAN, KNOUL
/ APPLICANT: ROGAN, PETER
/ TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATI
/ FILE REFERENCE: 30307
/ CURRENT APPLICATION NUMBER: US/09/573,080A
/ CURRENT FILING DATE: 2000-05-16
/ NUMBER OF SEQ ID NOS: 479
/ SOFTWARE: Patentin version 3.0
/ SEQ ID NO 7
/ LENGTH: 2710
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: repeat region
/ LOCATION: (1)..(2710)
/ OTHER INFORMATION: charlie3
/ NAME/KEY: misc_feature
/ OTHER INFORMATION: n is a, c, g or t
/ PUBLICATION INFORMATION:
/ AUTHORS: Jurka, J; Malichiewicz, J; Milosavljevic, A
/ TITLE: Prototypic sequences for human repetitive DNA
/ JOURNAL: Journal of Molecular Evolution
/ VOLUME: 35
/ ISSUE: 4
/ PAGES: 286-291
/ DATE: 1992-10-
/ DATABASE ACCESSION NUMBER: Database for repetitive elements (repbase)
/ DATABASE ENTRY DATE: 1996-01-26
US-09-573-080A-7

Query Match          60.7%; Score 18.2; DB 4; Length 2710;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 20; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
```

Qy 3 TGCTGTCAACCGTTCGTTGTT 27
 |||||
 Db 1209 TGCTGTTCACAGTTCGTTKGT 1185

RESULT 15
 US-09-949-016-200252/c
 ; Sequence 200252, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 200252
 ; LENGTH: 601
 ; TYPE: DNA
 ; ORGANISM: Human
 US-09-949-016-200252

Query Match 60.0%; Score 18; DB 4; Length 601;
 Best local Similarity 80.8%; Pred. No. 1.6e+02;
 Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 Qy 5 CTTGTTCACCGTTCGTTGTTGCA 30
 |||||
 Db 239 CTTGTAAACAGTTCGTTGCTCA 214

Search completed: September 15, 2005, 08:25:22
 Job time : 6.53786 secs

Sat Sep 17 09:06:48 2005

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 07:10:11 ; Search time 19.5379 Seconds
(without alignment)

10230.248 Million cell updates/sec

Title: US-10-713-381-2_COPY_1179_1208

Perfect score: 30
Sequence: 1 catgctgttcaacgcttgcgttcttcca 30Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7389322 seqs, 333185599 residues

Total number of hits satisfying chosen parameters: 14778644

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Published Applications NA:*

- 1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/prodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/prodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/prodata/2/pubpna/PCRTUS_PUBCOMB.seq:*
- 7: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/prodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/prodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/prodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/prodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/prodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/prodata/2/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/prodata/2/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/prodata/2/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/prodata/2/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/prodata/2/pubpna/US10G_PUBCOMB.seq:*
- 20: /cgn2_6/prodata/2/pubpna/US10H_PUBCOMB.seq:*
- 21: /cgn2_6/prodata/2/pubpna/US10I_PUBCOMB.seq:*
- 22: /cgn2_6/prodata/2/pubpna/US10_NEW_PUB.seq:*
- 23: /cgn2_6/prodata/2/pubpna/US11A_PUBCOMB.seq:*
- 24: /cgn2_6/prodata/2/pubpna/US11_NEW_PUB.seq:*
- 25: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq:*
- 26: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	100.0	30	US-10-713-381-4	Sequence 4, Appli
2	30	100.0	25	US-10-713-381-9	Sequence 9, Appli
3	30	100.0	1394	US-10-713-381-1	Sequence 1, Appli
4	30	100.0	1394	US-10-713-381-2	Sequence 2, Appli
5	24.2	80.7	158	US-10-713-381-3	Sequence 3, Appli
6	22	73.3	1304	US-10-001-857-90	Sequence 90, Appl
7	21	70.0	1980090	US-10-713-993-6615	Sequence 6615, Ap

C	8	21	70.0	1980090	21	US-10-741-600-17676	Sequence 17676, A
	9	20.4	68.0	1542	9	US-09-938-842A-1624	Sequence 1624, Ap
	10	20.2	68.0	1542	11	US-09-938-842A-1624	Sequence 1624, A
	11	20.2	67.3	2397	17	US-10-282-122A-31883	Sequence 31883, A
	12	20	66.7	2000	17	US-10-260-238-1785	Sequence 1785, A
	C	13	19.6	65.3	640	US-10-437-963-96639	Sequence 96639, A
	C	14	19.6	65.3	901	US-10-425-115-3534	Sequence 3534, Ap
	C	15	19.6	65.3	2000	US-10-260-228-1977	Sequence 1977, Ap
	C	16	19.4	64.7	2243	US-09-443-745-15	Sequence 15, Appl
	C	17	19.4	64.7	2256646	US-10-470-565-1	Sequence 1, Appli
	C	18	19	63.3	1041	US-10-425-115-100318	Sequence 100318, A
	C	19	63.3	1821	17	US-10-425-563-5	Sequence 5, Appli
	C	20	19	63.3	1821	US-10-972-789A-5	Sequence 5, Appli
	C	21	19	63.3	5671	US-09-764-872-729	Sequence 729, App
	C	22	19	63.3	198522	US-10-087-132-244	Sequence 244, App
	C	23	18.8	62.7	883	US-09-934-455-279	Sequence 279, App
	C	24	18.8	62.7	883	US-10-225-068A-183	Sequence 183, App
	C	25	18.8	62.7	883	US-10-225-068A-971	Sequence 971, App
	C	26	18.8	62.7	883	US-10-374-780A-381	Sequence 381, App
	C	27	18.8	62.7	883	US-10-225-068-183	Sequence 183, App
	C	28	18.8	62.7	883	US-10-225-068A-971	Sequence 971, App
	C	29	18.8	62.7	1083	US-09-815-242-7281	Sequence 7281, App
	C	30	18.8	62.7	1083	US-10-282-122A-22639	Sequence 22639, A
	C	31	18.8	62.7	1083	US-10-335-977-2194	Sequence 2194, Ap
	C	32	18.8	62.7	1125	US-10-335-977-2195	Sequence 2195, Ap
	C	33	18.8	62.7	1556	US-10-437-963-40048	Sequence 40048, A
	C	34	18.8	62.7	1899	US-10-472-928-2217	Sequence 2217, Ap
	C	35	18.8	62.7	1902	US-09-815-242-9220	Sequence 9220, Ap
	C	36	18.8	62.7	1902	US-10-282-122A-37788	Sequence 37788, A
	C	37	18.8	62.7	2000	US-10-260-238-2535	Sequence 2535, A
	C	38	18.8	62.7	2247	US-10-369-483-13483	Sequence 13483, A
	C	39	18.8	62.7	2247	US-10-108-260A-1041	Sequence 1041, App
	C	40	18.8	62.7	8788	US-10-756-149-64	Sequence 64, Appl
	C	41	18.8	62.7	9441	US-08-961-527-132	Sequence 132, App
	C	42	18.8	62.7	9441	US-10-158-844-132	Sequence 132, App
	C	43	18.8	62.7	2165598	US-10-472-928-4979	Sequence 4979, Ap
	C	44	18.6	62.0	375	US-09-974-300-3242	Sequence 3242, Ap
	C	45	18.4	61.3	174	US-09-814-353-5230	Sequence 5230, Ap

ALIGNMENTS

RESULT 1
US-10-713-381-4
; Sequence 4, Application US/10713381
; Publication No. US20040221331A1
GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-4

Query Match 100.0%; Score 30; DB 20; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.0065;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CATGCTTTCACCGCTTGTCTTTCACA 30

Db 1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
|||||

RESULT 2
US-10-713-381-9
; Sequence 9, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 255
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-9

Query Match 100.0%; Score 30; DB 20; Length 255;
Best Local Similarity 100.0%; Pred. No. 0.0086;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
|||||

Db 39 CATGCTTGTTCAACCGTTCGTTGTTCCA 68
|||||

RESULT 3
US-10-713-381-1
; Sequence 1, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-1

Query Match 100.0%; Score 30; DB 20; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
|||||

Db 1179 CATGCTTGTTCAACCGTTCGTTGTTCCA 1208
|||||

RESULT 4

US-10-713-381-2
; Sequence 2, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-2

Query Match 100.0%; Score 30; DB 20; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
|||||

Db 1179 CATGCTTGTTCAACCGTTCGTTGTTCCA 1208
|||||

RESULT 5
US-10-713-381-3
; Sequence 3, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 158
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-3

Query Match 80.7%; Score 24.2; DB 20; Length 158;
Best Local Similarity 89.7%; Pred. No. 2.2;
Matches 26; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CATGCTTGTTCAACCGTTCGTTGTTCC 29
|||||

Db 25 CATGCTTGTTCAACCGTTCGTTGTTCC 53
|||||

RESULT 6
US-10-001-857-90/c
; Sequence 90, Application US/10001857
; Publication No. US20020183500A1
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto

```
/ APPLICANT: Recipon, Hervé
/ APPLICANT: Chen, Sei-Yu
/ ORGANISM: Homo sapiens
/ APPLICANT: Sun, Yongming
/ APPLICANT: Liu, Chengshua
/ TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Pro
/ FILE REFERENCE: DEX-0273
/ CURRENT APPLICATION NUMBER: US/10/001,857
/ PRIOR APPLICATION NUMBER: 2001-11-20
/ PRIOR FILING DATE: 2000-11-20
/ NUMBER OF SEQ ID NOS: 208
/ SOFTWARE: PatentIn Version 3.1
/ SEQ ID NO 90
/ LENGTH: 1304
/ TYPE: DNA
/ ORGANISM: Homo sapien
US-10-001-857-90

Query Match      73.3%; Score 22; DB 13; Length 1304;
Best Local Similarity 83.3%; Pred. No. 25;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
DB      1155 CTTCTTGTTCCCGCTCTCTGTTTCCA 1126

RESULT 7
US-10-719-993-6815/c
/ Sequence 6815, Application US/10719993
/ Publication No. US2004025849A1
/ GENERAL INFORMATION:
/ APPLICANT: CARGILL, Michele et al.
/ TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
/ FILE REFERENCE: CL001496
/ CURRENT APPLICATION NUMBER: US/10/719,993
/ CURRENT FILING DATE: 2003-11-24
/ NUMBER OF SEQ ID NOS: 55342
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 6815
/ LENGTH: 1980090
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: (1)...(1980090)
/ OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-719-993-6815

Query Match      70.0%; Score 21; DB 20; Length 1980090;
Best Local Similarity 82.8%; Pred. No. 1.7e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      2 ATGCTTGTTCAACCGTTCGTTGTTCCA 30
DB      1948918 ATTCTTGCTGAACAGTTCGTTGTACCA 1948890

RESULT 8
US-10-741-600-17676/c
/ Sequence 17676, Application US/10741600
/ Publication No. US20050026169A1
/ GENERAL INFORMATION:
/ APPLICANT: CARGILL, Michele et al.
/ TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
/ FILE REFERENCE: CL001499
/ CURRENT APPLICATION NUMBER: US/10/741,600
/ CURRENT FILING DATE: 2003-12-22
/ NUMBER OF SEQ ID NOS: 73997
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 17676

/ LENGTH: 1980090
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: (1)...(1980090)
/ OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-741-600-17676

Query Match      70.0%; Score 21; DB 21; Length 1980090;
Best Local Similarity 82.8%; Pred. No. 1.7e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      2 ATGCTTGTTCAACCGTTCGTTGTTCCA 30
DB      1948918 ATTCTTGCTGAACAGTTCGTTGTACCA 1948890

RESULT 9
US-09-938-842A-1624
/ Sequence 1624, Application US/0993842A
/ Patent No. US20020160378A1
/ GENERAL INFORMATION:
/ APPLICANT: Harper, Jeff
/ APPLICANT: Kreps, Joel
/ APPLICANT: Wang, Xun
/ APPLICANT: Zhu, Tong
/ TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
/ FILE REFERENCE: SCRIP1300-3
/ CURRENT APPLICATION NUMBER: US/09/938,842A
/ CURRENT FILING DATE: 2001-08-24
/ PRIOR APPLICATION NUMBER: US 60/227,866
/ PRIOR FILING DATE: 2000-08-24
/ PRIOR APPLICATION NUMBER: US 60/264,647
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/300,111
/ PRIOR FILING DATE: 2001-06-22
/ NUMBER OF SEQ ID NOS: 5379
/ SEQ ID NO 1624
/ LENGTH: 1542
/ TYPE: DNA
/ ORGANISM: Arabidopsis thaliana
US-09-938-842A-1624

Query Match      68.0%; Score 20.4; DB 9; Length 1542;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
DB      237 CATGATGTTCAACCAATGTCGTTTCCA 266

RESULT 10
US-09-938-842A-1624
/ Sequence 1624, Application US/0993842A
/ Publication No. US20040009476A9
/ GENERAL INFORMATION:
/ APPLICANT: Harper, Jeff
/ APPLICANT: Kreps, Joel
/ APPLICANT: Wang, Xun
/ APPLICANT: Zhu, Tong
/ TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
/ FILE REFERENCE: SCRIP1300-3
/ CURRENT APPLICATION NUMBER: US/09/938,842A
/ CURRENT FILING DATE: 2001-08-24
/ PRIOR APPLICATION NUMBER: US 60/227,866
/ PRIOR FILING DATE: 2000-08-24
/ PRIOR APPLICATION NUMBER: US 60/264,647
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/300,111
```

;; PRIOR FILING DATE: 2001-06-22
;; NUMBER OF SEQ ID NOS: 5379
;; SEQ ID NO 1624
;; LENGTH: 1542
;; TYPE: DNA
;; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1624

Query Match 68.0%; Score 20.4; DB 11; Length 1542;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CATGCTGTTCACCGTTCGTTGTTCCA 30
Db 237 CATGATTGTTCAACCAATCGTCTTCCA 266

RESULT 11
US-10-282-122A-31883
; Sequence 31883, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITTA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31883
; LENGTH: 2397
; TYPE: DNA
; ORGANISM: Pseudomonas putida
US-10-282-122A-31883

Query Match 67.3%; Score 20.2; DB 17; Length 2397;
Best Local Similarity 88.0%; Pred. No. 1.6e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 CATGCTGTTCACCGTTCGTTG 25
|||||

Db 516 CATGCTGTTCACCGTTCGCTTG 540

RESULT 12
US-10-260-238-1785
; Sequence 1785, Application US/10260238
; Publication No. US20040016025A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, Paul R.
; APPLICANT: Moughamer, Todd G.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Katagiri, Puniyaki
; APPLICANT: Krep, Joel
; APPLICANT: Provart, Nicholas
; APPLICANT: Rieke, Darrell
; APPLICANT: Zhu, Tong

; TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION

; FILE REFERENCE: 60111-NP
; CURRENT APPLICATION NUMBER: US/10/260,238
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/325,448
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 6077
; SEQ ID NO 1785
; LENGTH: 2000
; TYPE: DNA

; ORGANISM: Oryza sativa
US-10-260-238-1785

Query Match 66.7%; Score 20; DB 17; Length 2000;
Best Local Similarity 82.1%; Pred. No. 1.9e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 TCCTTGTTCACCGTTCGTTGTTCCA 30
Db 1179 TACATGTTTACCGTTCGTTATTCAA 1206

RESULT 13
US-10-437-963-96639/C
; Sequence 96639, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Bouharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 96639
; LENGTH: 640
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_94716C.1
US-10-437-963-96639

Query Match 65.3%; Score 19.6; DB 19; Length 640;

Best Local Similarity 84.6%; Pred. No. 2.4e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 CTGTTCACCGTTCGTCCTGTTCGA 30
DB 199 CATGTTAACCGTTCGTCCTATTCAA 174

RESULT 14

US-10-425-115-3534/c
; Sequence 3534, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovacic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 3534
; LENGTH: 901
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_103225C.1
US-10-425-115-3534

Query Match 65.3%; Score 19.6; DB 20; Length 901;
Best Local Similarity 84.6%; Pred. No. 2.5e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CATGCTGTTCACCGTTCGTCCTGT 26
DB 33 CACGCTGTTCACCGTTCCTGCTGT 8

RESULT 15

US-10-260-238-1977/c
; Sequence 1977, Application US/10260238
; Publication No. US20040016025A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, Paul R.
; APPLICANT: Moushameh, Todd G.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Golf, Stephen A.
; APPLICANT: Katagiri, Fumiyuki
; APPLICANT: Kreps, Joel
; APPLICANT: Provart, Nicholas
; APPLICANT: Rieke, Darrell
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION
; FILE REFERENCE: 60111-NP
; CURRENT APPLICATION NUMBER: US/10/260,238
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/325,448
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 6077
; SEQ ID NO 1977
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Oryza sativa
US-10-260-238-1977

Query Match 65.3%; Score 19.6; DB 17; Length 2000;
Best Local Similarity 84.6%; Pred. No. 2.7e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 CTGTTCACCGTTCGTCCTGTTCGA 30
DB 291 CATGTTAACCGTTCGTCCTATTCAA 266

Search completed: September 15, 2005, 20:46:05
Job time : 26.5379 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 16:07:32 ; Search time 93.5658 Seconds
(without alignments)
12204.542 Million cell updates/sec

Title: US-10-713-381-2_COPY_1179_1208

Perfect score: 30
Sequence: 1 catgctgttcaacgcttcgctgttcca 30

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_ests1:*
2: gb_ests2:*
3: gb_nuc:*
4: gb_ests3:*
5: gb_ests4:*
6: gb_ests5:*
7: gb_ests6:*
8: gb_ests7:*
9: gb_ests8:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	30	100.0	687	9	CC656939 OGMDQ20TV
C 2	30	100.0	915	9	CG224225 OG1AG08TV
C 3	30	100.0	963	9	CC656933 OGMDQ20TV
C 4	22.2	74.0	308	9	CG700474 ZMMBC017
C 5	21.2	70.7	534	5	BU035739 CHU7N19.Y
C 6	21.2	70.7	626	5	BU030489 CHU15120
C 7	21.2	70.7	792	7	CN137013 OX1_54.E0
C 8	21	70.0	124	1	AA811551 0899E03.S
C 9	21	70.0	351	9	CR192459 Reverse s
C 10	21	70.0	571	9	CR064270 Reverse s
C 11	21	70.0	599	9	CR233964 Reverse s
C 12	21	70.0	769	9	CR188179 Reverse s
C 13	21	70.0	832	9	CC848693 NDL_2114.
C 14	21	70.0	835	9	CC860012 NDL_2113.
C 15	20.6	68.7	412	7	R12648
C 16	20.6	68.7	450	7	R12618
C 17	20.6	68.7	582	1	AU290776
C 18	20.6	68.7	681	8	AZ817790
C 19	20.6	68.7	1144	8	CC271930 CH261-112
C 20	20.4	68.0	353	8	AQ261553 CITBI-ET-
C 21	20.4	68.0	396	2	BE522930 M30D4STM
C 22	20.4	68.0	514	4	BG354154 00661_1ea
C 23	20.4	68.0	516	5	BP561510 BP561510
C 24	20.4	68.0	536	1	AV439701 AV439701

C 25	20.4	68.0	559	1	AV825937	AV825937
C 26	20.4	68.0	560	6	CA230882	CA230882
C 27	20.4	68.0	600	8	BH876120	BH876120
C 28	20.4	68.0	602	9	CL583422	CL583422
C 29	20.4	68.0	603	8	BZ694195	BZ694195
C 30	20.4	68.0	615	8	AQ453944	AQ453944
C 31	20.4	68.0	668	8	AO260121	AO260121
C 32	20.4	68.0	779	6	CD825116	CD825116
C 33	20.4	68.0	807	6	CD836362	CD836362
C 34	20.4	68.0	836	8	BH480110	BH480110
C 35	20.4	68.0	967	8	BZ684538	BZ684538
C 36	20.4	68.0	1205	8	B10570	B10570
C 37	20.2	67.3	431	7	R32995	R32995
C 38	20.2	67.3	565	6	CA681397	CA681397
C 39	20.2	67.3	729	8	BZ052097	BZ052097
C 40	20.2	67.3	791	9	CC836492	CC836492
C 41	20	66.7	170	9	DR5E15F	DR5E15F
C 42	20	66.7	227	4	B1130928	B1130928
C 43	20	66.7	329	4	B1127375	B1127375
C 44	20	66.7	329	7	CK098862	CK098862
C 45	20	66.7	368	4	B1128587	B1128587

ALIGNMENTS

RESULT 1
LOCUS OGMDQ20TV.ZM 0.7.1.5 KB Zea mays genomic clone ZMMEMa0554D15,
DEFINITION genomic survey sequence.
ACCESSION CC656939
VERSION CC656939.1 GI:32060231
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays

REFERENCE

AUTHORS Whiteclaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick,A., Fraser,C.M., Buddiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
TITLE Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Other GSSs: OGMDQ20TV
Contact: Cathy Whiteclaw
TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteclaw@tigr.org

Seq primer: TP
Class: sheared ends.

FEATURES

source Location/Qualifiers
1..687
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMEMa0554D15"
/clone_1lb="ZM_0.7.1.5_KB"
/note="vector: pBCSK-1 Site 1: HincII, 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match 100.0%; Score 30; DB 9; Length 687;
Best Local Similarity 100.0%; Pred. No. 0.02; Mismatches 0; Indels 0; Gaps 0;
Matches 30; Conservative 0;
1 CATGCTTGTCAACCGTTCGTTGTTCCA 30
|||||
397 CATGCTTGTCAACCGTTCGTTGTTCCA 368

RESULT 2
 CG224225 915 bp DNA linear GSS 22-AUG-2003
 LOCUS OGIAG08TV ZM 0.7 1.5 KB zea maye genomic clone ZMMBMA0716B15,
 DEFINITION genomic survey sequence.
 ACCESSION CG224225 GI:34124113
 VERSION CG224225.1 GI:34124113
 KEYWORDS GSS.
 SOURCE Zea maye
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 915)
 WhiteIaw,C.A., Quackenbush,J., Van Aken,S., Uteerback,T.,
 Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
 Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
 Consortium for Maize Genomics
 Unpublished (2002)
 Other_GSSs: OGIAG08TV
 Contact: Cathy WhiteIaw
 TIGR

TITLE
 JOURNAL
 COMMENT

FEATURES
 source
 1. 915
 Location/Qualifiers
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone_lib="ZMMBMA0716B15"
 /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
 methylation filtered genomic DNA library"

ORIGIN
 Query Match 100.0%; Score 30; DB 9; Length 915;
 Best Local Similarity 100.0%; Pred. No. 0.021;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
 |||||
 Db 558 CATGCTTGTTCAACCGTTCGTTGTTCCA 587

RESULT 3
 CC656933 963 bp DNA linear GSS 19-JUN-2003
 LOCUS OGMDQ20TM ZM 0.7 1.5 KB zea maye genomic clone ZMMBMA0554D15,
 DEFINITION genomic survey sequence.
 ACCESSION CC656933
 VERSION CC656933.1 GI:32060225
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 963)
 WhiteIaw,C.A., Quackenbush,J., Van Aken,S., Uteerback,T.,
 Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
 Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
 Consortium for Maize Genomics
 Unpublished (2002)
 Other_GSSs: OGMDQ20TV
 Contact: Cathy WhiteIaw
 TIGR

TITLE
 JOURNAL
 COMMENT

FEATURES
 source
 1. 963
 Location/Qualifiers
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone_lib="ZMMBMA0554D15"
 /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
 methylation filtered genomic DNA library"

ORIGIN
 Query Match 100.0%; Score 30; DB 9; Length 963;
 Best Local Similarity 100.0%; Pred. No. 0.021;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
 |||||
 Db 776 CATGCTTGTTCAACCGTTCGTTGTTCCA 805

RESULT 4
 CG700474 308 bp DNA linear GSS 15-OCT-2003
 LOCUS ZMMBMC0174H15f ZMMBMC (ECORI) Zea maye genomic clone ZMMBMC0174H15
 DEFINITION 5', genomic survey sequence.
 ACCESSION CG700474
 VERSION CG700474.1 GI:37688275
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 308)
 Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C.,
 Zohovetz,V., Fuks,G., Yu,Y., Wing,R. and Messing,J.
 Sequencing of the maize genome at PGR (2003c)
 Unpublished (2003)
 Contact: Bharti,A.K.
 Dr. Joachim Messing's lab
 The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
 University
 190 Frelinghuysen Road, Piscataway, NJ 08854, USA
 Tel: 732 445 3801
 Fax: 732 445 5735
 Email: bharti@waksman.rutgers.edu
 Seq primer: 17
 Class: BAC ends
 High quality sequence start: 57.
 Location/Qualifiers
 1. 308
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /cultivar="B73"
 /db_xref="taxon:4577"
 /clone_lib="ZMMBMC0174H15"
 /lab_host="E. coli DH10B"
 /clone_lib="ZMMBMC (ECORI)"
 /note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI"

ORIGIN
 Query Match 74.0%; Score 22.2; DB 9; Length 308;
 Best Local Similarity 88.9%; Pred. No. 49;
 Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CATGCTTGTTCAACCGTTCGTTGTT 27

[illegible]

ACCESSION	BU030489
VERSION	BU030489.1
KEYWORDS	GI:22466009
SOURCE	EST
ORGANISM	Helianthus annuus (common sunflower)
REFERENCE	Helianthus annuus Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; campanulids; Asterales; Asterales; Asteraceae; Asteroidae; Heliantheae; Helianthus.
AUTHORS	1 (bases 1 to 626) Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L., Lin, H., van Damme, M., Lavelle, D., Chevalier, P., Ziegler, J., Ellison, P., Kolman, J., Stabaugh, M.S., Livingston, K., Zhou, Y., Lai, Z., Church, S., Jackson, L., and Bradford, K. Lecture and Sunflower ESTs from the Compositae Genome Project http://compgenomics.ucdavis.edu/ Unpublished (2002)
JOURNAL	Contact: Alexander Kozik [R.W.Michelmore] Department of Vegetable Crops, R.W.Michelmore lab University of California at Davis (UCD) Asmundo Hall, UCD, Davis, CA 95616, USA Tel: 1-(530)-742-1742 Fax: 1-(530)-752-9659 Email: akozik@ucdavis.org [michelmore@vegmail.ucdavis.edu] belongs to contig OH_Ca_Conc193046, see http://cgdb.ucdavis.edu/ for details.
COMMENT	Plate: OHJ15 row: I column: 20. Location/Qualifiers 1..626 /organism="Helianthus annuus" /mol_type="mRNA" /cultivar="RHA280" /db_xref="taxon:4232" /clone="OHJ15120" /lab_host="E. coli" /clone_lib="OH_ERGHU sunflower RHA280" /note="Vector: pBRC2VAST11ab: The library was constructed from 11 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgdb.ucdavis.edu/ TAG_Lib=OH_ERGHU sunflower RHA280 TAG_SEQ=CGAATCGG3"
ORIGIN	
Query Match	70.7%; Score 21.2; DB 5; Length 626;
Best Local Similarity	88.5%; Pred. No. 1.5e+02;
Matches	23; Conservative 0; Mismatches 3; Indels 0; Gaps 0.
Dy	2 ATGCTGTTCACCGTTCGTTGTT 27 104 ATGCTTCTCAAGCGTTCGTTT 79
RESULT 7	
CN137013	792 bp mRNA linear EST 01-APR-2004
LOCUS	OX1_54_E01.A002 Oxidatively-stressed leaves and roots Sorghum
DEFINITION	bicolor cDNA clone OX1_54_E01.A002 5', mRNA sequence.
ACCESSION	CN137013
VERSION	CN137013.1
KEYWORDS	GI:45970259
SOURCE	Sorghum bicolor (sorghum)
ORGANISM	Sorghum bicolor Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Sorghum. 1 (bases 1 to 792)

ORIGIN
Query Match 70.0%; Score 21; DB 9; Length 351;
Best Local Similarity 82.8%; Pred. No. 1.7e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 2 ATGCTGTTCACCGCTTCGTCTGTTC 30
DB 307 ATGCTGTTCACCGCTTCGTATGTTCA 279

RESULT 10
CR064270 571 bp DNA linear GSS 05-JUL-2004
LOCUS CR064270
DEFINITION Reverse strand read from insert in 5'HPRT insertion targeting and
chromosome engineering clone MHPN190119, genomic survey sequence.
ACCESSION CR064270
VERSION CR064270.1 GI:49797742
KEYWORDS GSS; genome survey sequence; MICER.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
Rogers,J., and Bradley,A.
TITLE Direct Submision
JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. <http://www.sanger.ac.uk/MICER>
Location/Qualifiers
FEATURES
source 1..571
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="MHPN190119"
/clone_1lb="MHPN"

ORIGIN
Query Match 70.0%; Score 21; DB 9; Length 571;
Best Local Similarity 82.8%; Pred. No. 1.8e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 2 ATGCTGTTCACCGCTTCGTCTGTTC 30
DB 309 ATGCTGTTCACCGCTTCGTATGTTCA 281

RESULT 11
CR233964 599 bp DNA linear GSS 06-JUL-2004
LOCUS CR233964
DEFINITION Reverse strand read from insert in 5'HPRT insertion targeting and
chromosome engineering clone MHPN14h08, genomic survey sequence.
ACCESSION CR233964
VERSION CR233964.1 GI:50012813
KEYWORDS GSS; genome survey sequence; MICER.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 599)
Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
Rogers,J., and Bradley,A.
TITLE Direct Submision
JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. <http://www.sanger.ac.uk/MICER>
Location/Qualifiers
FEATURES
source 1..599
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="MHPN14h08"

ORIGIN
Query Match 70.0%; Score 21; DB 9; Length 599;
Best Local Similarity 82.8%; Pred. No. 1.8e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 2 ATGCTGTTCACCGCTTCGTCTGTTC 30
DB 310 ATGCTGTTCACCGCTTCGTATGTTCA 282

RESULT 12
CR188179 769 bp DNA linear GSS 06-JUL-2004
LOCUS CR188179
DEFINITION Reverse strand read from insert in 5'HPRT insertion targeting and
chromosome engineering clone MHPN234h03, genomic survey sequence.
ACCESSION CR188179
VERSION CR188179.1 GI:49967028
KEYWORDS GSS; genome survey sequence; MICER.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 769)
Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
Rogers,J., and Bradley,A.
TITLE Direct Submision
JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. <http://www.sanger.ac.uk/MICER>
Location/Qualifiers
FEATURES
source 1..769
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="MHPN234h03"
/clone_1lb="MHPN"

ORIGIN
Query Match 70.0%; Score 21; DB 9; Length 769;
Best Local Similarity 82.8%; Pred. No. 1.9e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 2 ATGCTGTTCACCGCTTCGTCTGTTC 30
DB 307 ATGCTGTTCACCGCTTCGTATGTTCA 279

RESULT 13
CC848693 832 bp DNA linear GSS 24-JUL-2003
LOCUS CC848693
DEFINITION ND1.2114 T7 Notre Dame Liverpool Aedes aegypti genomic clone
Notredame Liverpool-2114, genomic survey sequence.
ACCESSION CC848693
VERSION CC848693.1 GI:33197857
KEYWORDS GSS.
SOURCE Aedes aegypti (yellow fever mosquito)
ORGANISM Aedes aegypti

REFERENCE
AUTHORS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Aedes;
Stegomyia.
1 (bases 1 to 832)
Loftus,B., Shetty,J., Knudson,D. and Severson,D.
TITLE BAC end sequencing of Aedes aegypti
JOURNAL Unpublished (2003)
COMMENT Other_GSSs: ND1.2114.SP6
Contact: Brendan Loftus
Department of Eukaryotic Genomics
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-3543
Fax: 301-838-0208

Email: enta@tigr.org
Library was provided by David Severson
Seq primer: T7
Class: BAC ends.

FEATURES

source

Location/Qualifiers

1..832
/organism="Aedes aegypti"
/mol_type="genomic DNA"
/strain="Liverpool"
/db_xref="taxon:7159"
/clone="Notre Dame Liverpool-2114"
/note="Vector: pECBAC1, Site 1: Hind III; The library was prepared from whole body tissue of newly hatched L1 larvae by David Severson at the University of Notre Dame and Hongbin Zhang"

ORIGIN

Query Match 70.0%; Score 21; DB 9; Length 832;

Best Local Similarity 82.8%; Pred. No. 1.9e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OR

2 ATGCTTTGTCACCGTTCCTTTGTTCCA 30
|||||
242 ATCCTTTGTCACCATTTGCTTGATTTCA 214

DB

242 ATCCTTTGTCACCATTTGCTTGATTTCA 214

RESULT 14

CC860012/c

LOCUS CC860012 835 bp DNA linear GSS 24-JUL-2003
DEFINITION NDL.2113.T7 Notre Dame Liverpool Aedes aegypti genomic clone
Notre Dame Liverpool-2113, genomic survey sequence.

ACCESSION

CC860012

VERSION

CC860012.1 GI:33220022

KEYWORDS

GSS.

SOURCE

Aedes aegypti

ORGANISM

Aedes aegypti

REFERENCE

1 (bases 1 to 835)

AUTHORS

Loftus,B., Shetty,J., Knudson,D. and Severson,D.

TITLE

BAC end sequencing of Aedes aegypti

JOURNAL

Unpublished (2003)

COMMENT

Other GSS: NDL.2113.SP6

CONTACT

Department of Eukaryotic Genomics

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel:

301-838-3543

Fax:

301-838-0208

Email:

enta@tigr.org

Library was provided by David Severson

Seq primer: T7

Classes: BAC ends.

Location/Qualifiers

1..835

/organism="Aedes aegypti"

/mol_type="genomic DNA"

/strain="Liverpool"

/db_xref="taxon:7159"

/clone="Notre Dame Liverpool-2113"

/note="Vector: pECBAC1, Site 1: Hind III; The library was prepared from whole body tissue of newly hatched L1 larvae by David Severson at the University of Notre Dame and Hongbin Zhang"

ORIGIN

Query Match 70.0%; Score 21; DB 9; Length 835;

Best Local Similarity 82.8%; Pred. No. 1.9e+02;

Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OR 2 ATGCTTTGTCACCGTTCCTTTGTTCCA 30
|||||
242 ATCCTTTGTCACCATTTGCTTGATTTCA 214

FEATURES

source

1..412
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:398607"
/db_xref="taxon:9606"
/clone="IMAGE:26260"
/sex="female"
/dev_stage="73 days post natal"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares infant brain INIB"
/note="Organ: whole brain; Vector: Lambda BAC; Site 1: Not I; Site 2: Hind III; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5',
ACTGGAAGAAATTCGCGCCGACAGAAATTTTCTTTTCTTTT 3'];
double-stranded cDNA was ligated to Hind III adaptors
(Pharmacia), digested with Not I and directionally cloned
into the Not I and Hind III sites of the lambda BAC vector.
Library went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match

Best Local Similarity

Matches

23; Conservative

0; Mismatches

4; Indels

0; Gaps

0;

Search completed: September 16, 2005, 08:08:43

Job time: 93.5658 secs

RESULT 1	AX224399	40 bp	DNA	linear	PAT 10-SEP-2001
LOCUS	AX224399				
DEFINITION	Sequence 6 from Patent WO0160997.				
ACCESSION	AX224399				
VERSION	AX224399.1	GI:15554641			
KEYWORDS					
SOURCE					
ORGANISM					
	Zea mays				
	Zea mays				
	Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.				
REFERENCE					
AUTHORS	1 Albertsen,M.C., Fox,T.W., Garmaat,C.W., Huffman,G. and Kendall,T.L.				
TITLE	Male tissue-preferred regulatory region and method of using same				
JOURNAL	Patent: WO 0160997-A 6 23-AUG-2001; PIONEER HI-BRED INTERNATIONAL, INC. (US)				
FEATURES					
source					
	1..40				
	/organism="Zea mays"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:4577"				
ORIGIN					
		100.0%;	Score 40;	DB 6;	Length 40;
Query Match		Best Local Similarity 100.0%;	Pred. No. 4.2e-05;		
Matches	40;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1 AGGATACCTACTCCCAACAATCCATCTTACTCATGCAAC 40				
Db	1 AGGATACCTACTCCCAACAATCCATCTTACTCATGCAAC 40				
RESULT 2	AX224398				
LOCUS	AX224398				
DEFINITION	Sequence 5 from Patent WO0160997.				
ACCESSION	AX224398				
VERSION	AX224398.1	GI:15554640			
KEYWORDS					
SOURCE					
ORGANISM					
	Zea mays				
	Zea mays				

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
AUTHORS
TITLE
JOURNAL
1
Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
Patent: WO 0160997-A 5 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)

FEATURES
source
1..50
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN

Query Match 100.0%; Score 40; DB 6; Length 50;
Best Local Similarity 100.0%; Pred. No. 4.1e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGATACCTACTCCCAACATCCATCTTACTCATGCAAC 40
|||||
11 AGGATACCTACTCCCAACATCCATCTTACTCATGCAAC 50

Db 11 AGGATACCTACTCCCAACATCCATCTTACTCATGCAAC 50

RESULT 3
AX224396 158 bp DNA linear PAT 10-SEP-2001
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
AUTHORS
TITLE
JOURNAL
1
Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
Patent: WO 0160997-A 3 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)

FEATURES
source
1..158
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN

Query Match 100.0%; Score 40; DB 6; Length 158;
Best Local Similarity 100.0%; Pred. No. 3.6e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGATACCTACTCCCAACATCCATCTTACTCATGCAAC 40
|||||
86 AGGATACCTACTCCCAACATCCATCTTACTCATGCAAC 125

Db 86 AGGATACCTACTCCCAACATCCATCTTACTCATGCAAC 125

RESULT 4
AX224402 255 bp DNA linear PAT 10-SEP-2001
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
AUTHORS
TITLE
JOURNAL
1
Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
Patent: WO 0160997-A 9 23-AUG-2001;

PIONEER HI-BRED INTERNATIONAL, INC. (US)

FEATURES
source
1..255
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN

Query Match 100.0%; Score 40; DB 6; Length 255;
Best Local Similarity 100.0%; Pred. No. 3.4e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGATACCTACTCCCAACATCCATCTTACTCATGCAAC 40
|||||
99 AGGATACCTACTCCCAACATCCATCTTACTCATGCAAC 138

Db 99 AGGATACCTACTCCCAACATCCATCTTACTCATGCAAC 138

RESULT 5
AX224394 1394 bp DNA linear PAT 10-SEP-2001
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
AUTHORS
TITLE
JOURNAL
1
Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
Patent: WO 0160997-A 1 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)

FEATURES
source
1..1394
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN

Query Match 100.0%; Score 40; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 2.8e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGATACCTACTCCCAACATCCATCTTACTCATGCAAC 40
|||||
1239 AGGATACCTACTCCCAACATCCATCTTACTCATGCAAC 1278

Db 1239 AGGATACCTACTCCCAACATCCATCTTACTCATGCAAC 1278

RESULT 6
AX224395 1394 bp DNA linear PAT 10-SEP-2001
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
AUTHORS
TITLE
JOURNAL
1
Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
Patent: WO 0160997-A 2 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)

FEATURES
source
1..1394
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN

Query Match 100.0%; Score 40; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 2.8e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGATACCTACTCCCAACAATCCATCTTACTCATGCAAC 40
|||||
Db 1239 AGGATACCTACTCCCAACAATCCATCTTACTCATGCAAC 1278

RESULT 7
BD062176 1394 bp DNA linear PAT 27-AUG-2002
LOCUS Male tissue-preferred regulatory region and method of using same.
DEFINITION BD062176.1 GI:22607781
ACCESSION BD062176.1
VERSION JP 2001520523-A/1.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1394)
Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G.A. and
Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
PATENT: JP 2001520523-A 1 30-OCT-2001;
PIONEER HI BRED INTERNATIONAL INC
PN JP 2001520523-A/2

COMMENT

TITLE Male tissue-preferred regulatory region and method of using same
JOURNAL PIONEER HI BRED INTERNATIONAL INC
AUTHORS Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G.A. and
Kendall,T.L.
REFERENCE 1 (bases 1 to 1394)
Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G.A. and
Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
PATENT: JP 2001520523-A 1 30-OCT-2001;
PIONEER HI BRED INTERNATIONAL INC
PN JP 2001520523-A/1

FEATURES
source 1. .1394
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Key Location/Qualifiers.
FH

Query Match 100.0%; Score 40; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 2.8e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGATACCTACTCCCAACAATCCATCTTACTCATGCAAC 40
|||||
Db 1239 AGGATACCTACTCCCAACAATCCATCTTACTCATGCAAC 1278

RESULT 8
BD062177 1394 bp DNA linear PAT 27-AUG-2002
LOCUS Male tissue-preferred regulatory region and method of using same.
DEFINITION BD062177.1 GI:22607782
ACCESSION BD062177.1
VERSION JP 2001520523-A/2.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1394)
Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G.A. and
Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
PATENT: JP 2001520523-A 2 30-OCT-2001;
PIONEER HI BRED INTERNATIONAL INC
PN JP 2001520523-A/2

COMMENT

BD 30-OCT-2001
PF 19-JUN-1998 JP 1999504910
PR 23-JUN-1997 US 08/880499
PI MAC C ALBERTSEN,TIMOTHY W FOX,CARL W GARNAA,T,GARY A HUFFMAN,
PI TIMMY L KENDALL
PC C12N15/82,C12N15/29,C12N9/24,C12N9/22,C12N9/10,C12N9/00 PC
C07K14/34,C1201/68,
PC A01H5/00
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers.
source 1. .1394
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Key Location/Qualifiers.
FH

Query Match 100.0%; Score 40; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 2.8e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGATACCTACTCCCAACAATCCATCTTACTCATGCAAC 40
|||||
Db 1239 AGGATACCTACTCCCAACAATCCATCTTACTCATGCAAC 1278

RESULT 9
AF360356 3343 bp DNA linear PLN 12-MAY-2001
LOCUS Zea mays male fertility protein (Ms45) gene, complete cds.
DEFINITION AF360356
ACCESSION AF360356
VERSION AF360356.1 GI:14028756
KEYWORDS Zea mays
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 3343)
Fox,T.W., Trimmell,M.R. and Albertsen,M.C.
Cloning of Ms45, a gene required for male fertility from Zea mays
Unpublished
2 (bases 1 to 3343)
Fox,T.W., Trimmell,M.R. and Albertsen,M.C.
Direct Submission
Submitted (13-MAR-2001) Trait and Technology Development, Pioneer
Hi-Bred Intl. Inc., 7300 N.W. 62nd Ave., P.O. Box 1004, Johnston,
IA 50131-1004, USA
LA 50131-1004, USA
Location/Qualifiers
source 1. .3343
/organism="Zea mays"
/mol_type="genomic DNA"
/culivar="B73"
/db_xref="taxon:9606"
/chromosome="9L"
<1392..>2942
/gene="Ms45"
/join<1392..1768,1898..2182,2280..2447,2534..>2942)
/gene="Ms45"
/product="male fertility protein"
/protein_id="AAK52489.1"
/db_xref="GI:14028757"
/translation="MEKRNLOWRRGRDGIVOYPLHFFAALLALIVADPFGLSPLAEV
DYRPAVKEHIAPIYGEVWGSMPBDNARSRLRGRLFEVGEVGESEIFDLOGRPVYGLA
DGRVVRWNGEAGMETPAVMNPDMSGEVACVNSTTRKHQKEFCGRPIGLRPHGE
TGEIVVADAYGVGLAVVGOSGVASVAREADGDPTRFANDLDVHRNGSVFPDTSMRY
SKKHNLNLLIEGEGRLRLRYDPEPSTGVHVLKGI.VPENGQDISDHOELLPSPTTNC
RMRWYWGSPRAGEVVEVPANLDPFPDNRVNSRGQFVWAIIDCCTPDAOEVPKRPWLR

ORIGIN
TUYFKPLSLKLVITWKAARMHTVLLDDEGRVVELEDREGEVWKLVSSEVSEVGRK
LWIGTVAHHNHIATIPYPLED"

Query Match 100.0%; Score 40; DB 8; Length 3343;
Best Local Similarity 100.0%; Pred. No. 2,6e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AGGATACCTACTCCGAACATCATCTTACTCATGCAC 40
DB 1239 AGGATACCTACTCCGAACATCATCTTACTCATGCAC 1278

RESULT 10 AC135206 137327 bp DNA linear PLN 16-APR-2003
LOCUS AC135206/c Oryza sativa (japonica cultivar-group) chromosome 3 clone
DEFINITION OJ1041F02, complete sequence.
ACCESSION AC135206
VERSION AC135206.3 GI:27596977
KEYWORDS HTG.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Eriocaridaceae; Oryzaceae; Oryza.
1 (bases 1 to 137327)
Wing,R.A., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Currie,J. and
Collura,K.

REFERENCE AUTHORS
TITLE Rice Genomic Sequence
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 137327)
Wing,R.A., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Saski,C.,
Currie,J., Collura,K. and Thompson,S.
DIRECT SUBMISSION
SUBMITTED (09-OCT-2002) Arizona Genomics Institute, University of
Arizona, 303 Forbes, Tucson, AZ 85721, USA
3 (bases 1 to 137327)
Wing,R.A., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Saski,C.,
Currie,J., Collura,K. and Thompson,S.
DIRECT SUBMISSION
SUBMITTED (06-NOV-2002) Arizona Genomics Institute, University of
Arizona, 303 Forbes, Tucson, AZ 85721, USA
4 (bases 1 to 137327)
Wing,R.A., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Currie,J.,
Collura,K. and Thompson,S.

REFERENCE AUTHORS
TITLE Direct Submission
JOURNAL Submitted (11-JAN-2003) Arizona Genomics Institute, University of
Arizona, 303 Forbes, Tucson, AZ 85721, USA
5 (bases 1 to 137327)
Wing,R.A., Yu,Y., Soderlund,C., Chen,M., Kim,H., Rambo,T.,
Saski,C., Henry,D., Thompson,S., Simmons,J., Thurmond,S.K. and
Sun,S.

REFERENCE AUTHORS
TITLE Direct Submission
JOURNAL Submitted (16-APR-2003) Clemson University Genomics Institute,
Clemson University, 100 Jordan Hall, Clemson, SC 29634, USA
On Jan 11, 2003 this sequence version replaced gi:24635891.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., paired quality
>=30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by more than
one plasmid subclone; and the assembly was confirmed by
restriction digest. There are transposons located at area
29427-32874 and at area 19945-33212. Region 70511-70536 is covered
by a single subclone. Region 18678-18673 is a double stranded
single subclone. Areas 69957-70761, 89836-90416, 98951-99009,
105316-105607, and 133052-133152 are covered by Monsanto only. The
nucleotide sequence of this BAC clone was generated by combining
Syngenta, Monsanto and Arizona Genomics Institute sequencing data.

REFERENCE AUTHORS
TITLE Direct Submission
JOURNAL Submitted (16-APR-2003) Clemson University Genomics Institute,
Clemson University, 100 Jordan Hall, Clemson, SC 29634, USA
On Jan 11, 2003 this sequence version replaced gi:24635891.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., paired quality
>=30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by more than
one plasmid subclone; and the assembly was confirmed by
restriction digest. There are transposons located at area
29427-32874 and at area 19945-33212. Region 70511-70536 is covered
by a single subclone. Region 18678-18673 is a double stranded
single subclone. Areas 69957-70761, 89836-90416, 98951-99009,
105316-105607, and 133052-133152 are covered by Monsanto only. The
nucleotide sequence of this BAC clone was generated by combining
Syngenta, Monsanto and Arizona Genomics Institute sequencing data.

FEATURES
source 1.137327
/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="genomic DNA"
/db_xref="taxon:39947"
/chromosome="3"
/clone="OJ1041F02"
/note="(japonica cultivar-group)"
complement(638..9495)
/gene="OJ1041F02.1"

CDS
/note="unknown protein"
complement(join(638..838,1141..1276,3132..3244,3326..3421,
3513..3727,3812..3934,4024..4054,4157..4216,5088..5243,
5322..5411,5484..5699,5808..5942,6035..6538,6622..6837,
7548..7593,7679..7815,8196..8328,9221..9495))
/gene="OJ1041F02.1"

/codon_start=1
/product="unknown protein"
/protein_id="AAP06844.1"
/db_xref="GI:29893590"

/translation="MAQSLLELLIQFLMPNDARAEQIRRLARDPQVPAVYHL
RTAKTPNROLAAVLIRKKTISHMPLPPHAKASLQALIDSITDHSILYRABANV
VSIKXAVPAGEMPBELPFIQCSQSPQEDHREVALIFSLTETIGTTPQSHNDL
QPIILKICLQDASSSRVRIALKAQSPFIEVNDGVDVYKIFRDPISLIANSROCLAN
GEBDVASTAFETPDELISPPALUGDSRSTIVQPSLEVCNDELEINROAQIOLISM
LVKFKASFLKHKLVITPIQVWCPLTETAEDESDLAADRSAAEVINDTAIINPRH
VPPVLEPAAVSFRHINPKYREAATSVLSVSEBGCERLKKLQEDCLRVVLEALDOE
QVNRGAASFALGOFAHEHQPEILSHYSEVLICINALBDDPSBEVSEKSYVALAFCEQ
GMNLIPLDPLMCLVMSLQSPRNLOETGMSAIGSAVAAAEQAFMEYAEKYLEMMK
GMENTIENDELCARABATEVGIYMAVAGRAMETILPFIETASISGVLDVSELEY
THGFSNVAETLIDDSFAOYLPHVVPAASSCNLDGSAVDIDDAISVNGSCVSDD
DVNDEPRVRIISWLTITIDITAIIRAIPPAADVLEKODILDTWMIYIKTRWEDD
KEVAQACTSLADIVRDCGFALIEPYITRLADATILIRQSCCQVQSDDEGDDID
HDEVLMADVSLDLPFAFKVMSYEDPITFKLDSIMKPAKBPPODKTMVATLAEV
AQGMGAPISAVYDKIMPLVLEKELASSEATNRNNAFCVCEKNGGAALAKYGDILH
GLHRLFADESEDDAVRDNAAGAIARMIWVQOSIPLNOLIPLVPIVINAQVAVSPN
ESDEVKTVAKAVSHLISVGOQNPILSALPRAHANLASPARNR"

12372..15503
/gene="OJ1041F02.2"
/note="unknown protein"
join(12372..12757,12858..13025,13130..13295,13716..13826,
13982..14130,14224..14354,14459..14558,15197..15503)
/gene="OJ1041F02.2"

/codon_start=1
/product="unknown protein"
/protein_id="AAP06845.1"
/db_xref="GI:29893591"

/translation="MPQGGECRHGMLVGEEAAYGGGEPEFYVPLKRLSVDSKAS
TAPRICWECQGBAGDITCDIYAAPLRSCGSAKAMPPEAPLPFWMTPEPPRQSDGAS
EARRGEAIRKGRHSYSMLNLIQAGISISVSKSALPFOKLAASDPREKWTFRFP
EGSKTPPHHSVDFRWKDYCPAVFRPHKLFGVDPAEYMLAICGDTLRELASPGKSG
SCFFITQDDRMIKTVKKSSEVKVILRMIRSYEHVROYKSTLTFEYTHCIGKQCGT
KLHRRFDLKGSSHGRTIDTERKIDETTLQDQVAFRLQRFVEEGLMIOQMDCT
PLENGIMDYSLLGVHFRNDYSVKISGIIHIAPEKSTGRKSPFEGEGLSCFELFEV
SGCKDRDILDSRKPEIOLGIMPPQAQAESSKSIIDNPIINRHLFITPSSGSCDYVL
FGGIIIDLIQDYDITKLEHAYKSQVNDYIISANDPKLYSRFQDFIRRVFIKEQ"

16734..16878
/note="putative MITE, Galjin/Galjin-like"
19263..19454
/note="putative transposon, CACTA, En/Spm-like"
21819..24660
/gene="OJ1041F02.3"

CDS
/note="hypothetical protein"
join(12819..22007,22096..22206,23772..23996,24523..24660)
/gene="OJ1041F02.3"

/codon_start=1
/product="hypothetical protein"
/protein_id="AAP06846.1"
/db_xref="GI:29893592"

/translation="WATSDSGSGGEGQRLLSIPKEGRIIAPTRPDGTLRKAIRI
RAGVPOEVAIVIOSKGAQKRSKPDVPGYDPAIDAPKTKAARNRRKRRQOAS
TTNDGKGLHIEDAGETNDKQDVSTTKOISGIAEVLVAATSSDARDNSKSES
SAPDIDKIRALKKIRALEAAYGVGDPENLKPQLEKSKKIEGMEBELKLENNSSPA
AS"

complement(26563..30575)
/gene="OJ1041F02.4"

/note="unknown protein"
complement (join126563. .26772, 27082. .27288, 27430. .27537,
27881. .27940, 28051. .28126, 28327. .28464, 28616. .28678,
28790. .28874, 29560. .30575))
/gene="OJ1041F02.4"
/codon_start=1
/product="unknown protein"
/protein_id="AAP06847.1"
/db_xref="GI:29893593"
/translation="MGPPAPSPSSSGSSRRRLERRNAKHIGVDASNFCAVP
QSPAPASGSPSLACSPASLIDLSFPIGSGGDCRDVQLCSLGLSGVDVAV
PVADWEAHKAGRSSFSSTASTPKRREPPARDSPVREVAEPEPSLPAPAPLPA
AKETPSVAIEAPPLRVDPMEPARDVAKSGEGGICKVREPVYLKPEWMPVPA
VCVEFTWDLRSPAPEDSHAPASRCGDSACODAGEEDAAVLTRELRGE
TSREFTSSLSLTNDERTSTTESPFYISNGRRRKTRSRMLTSGSGFTYF
EGISDEGVFAVKEVCLCDGSAQOCTFQLEQELALISQHEHNVQYGTKESEK
LYIFLEVTGSLASLVQKRLDTHVATRIINGLTLYLHRNIVHRIKCANILY
HANGSVLADDFGLAKEITKENVLKSCGYVYMAPEVYNKTYGSEADLWSGCTVL
EMLTROPYGLMETQALYRIGGEPAIENGCRADRPISOCVKNPDORSAKL
LEHPVNRSMRSIRSMRTSRNSITVILIMSMGCRTRIRYEALISTQSTPTIA
PHLLSHSLRGDMFSORLLFSPENRLNTHSA"

repeat_region
/note="putative MITE, MITE-adh, type D-like"
27679. .27832
complement (31668. .31815)
complement (35049. .49855)
/gene="OJ1041F02.5"
/note="hypothetical protein"
complement (join135049. .35184, 35866. .36033, 36922. .37205,
37280. .38523, 40038. .40104, 40196. .40287, 40406. .40617,
41678. .41760, 42229. .42259, 42907. .43088, 43694. .43735,
44229. .44426, 44622. .44676, 45877. .45992, 46149. .46313,
46447. .46620, 47271. .47470, 47915. .48019, 48200. .48263,
49589. .49855)
/gene="OJ1041F02.5"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAP06848.1"
/db_xref="GI:29893594"
/translation="MGSDADPEAPPTPTSPSPAPATPSADNGNRLRRCYQSLTS
WGPPKAGGGGAGAGLPLLAGDGTPEPKRGRPRSEAGKSSNRETTGLSD
SKDEVILVSEPKQKRGKNGGALKVPRKCKALESTGDSNCOOLRSQDRA
VLPOKSLTVDIDLVTGPRSEAPKNDVNDALDNEDKPOLIVDRSEANIAEENRSL
SGKKMPFASRKIHGAGODILNVEDMDSLCAFEPDQDLPVHLYELVTMELH
MSNKLILADSKPLCTSTRENSAHDRECHLANPHDKONKRFSSQDVIDUDELL
ASSSCFASLIFESKQHERVQHELEVTPEKCOQJANLMDTRETRELAQVCSNKHAF
LNEALGWBERGHRNKONITNSNGRSQDSDIDYSDADYNVLLITGPVCG
KSAVPAFCAREGFNVIELNPKTVFPPPHNKLQCTDLDPDKLHSSHSIKYI
PTRVPMNRNSIISISAHSSSLVILKHLVNTSMRNGAYVROKFEETKSHLEK
MSOEIIIGLPISSLDPASGTPGTAHYKQVINKTILIFEDVDVPEDEGRTSTIKM
VETTKPPIITSNKDPPLPHLAQLVLDFTYSSABLSHDMICKSBVEITVPOQ
KATIDAFGLKNCSCPLDDADVASTVPRIMPDPCKSETTYMEIDKTVIAE
OKKQMEVSEFEGLQIMPLTKGASAGTRPKSKLKGSDACNAPSPCKNDL
DFHSDPDIPLPSNORMNRNGVVLPAESDDDLADNAADAFTVQEGELPQSSDL
PCYAGISINIVSEVFOOSVPHLREVIENQLCPEPSSRFBFASRONGLEENM
SVASIOICDTPFOGITSCEPSESFMVGTGSASISDPLISLIVNSILNRSYTYA
SVALDPTKVENQNTDKQCKMEDEVGECYAYELABRNDASCTIGYQIMDCS
RAASVWLISGKNKNSCKEAVQDVTNRNLKQCHVLPCKMNRHSVSGALKVRSRD
LISDELMLISCHPFSNDISDPLRTYTESDGSYSKOLEMGSIVYOHGICFLQSO
ATDEGVDVLOELIFSGTTTSLGKFVSSGISGDSGNSHXYTSCISKREOA
RLAEVLLPVVPPKLSQSLRGPAFVYLSMSOILENNOLSECKASQKORCRORRH
YLSGSLISAEIIGLAAQSTFSDRRESSTIEQALS"

repeat_region
39012. .39207
/note="putative MITE, MITE-adh, type M-like"
39681. .39879
/note="putative MITE, MITE-adh, type G-like"
51474. .51836
/note="putative MITE, Caatcaway-like"
complement (59793. .62196)
/gene="OJ1041F02.6"

Query Match 76.0%; Score 30.4; DB 8; Length 137327;
Best Local Similarity 85.0%; Pred. No. 0.12;
Matches 34; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAACCAATCCATCTTACTATGTCAC 40
DB 119478 AGAACAACCTACTCCCAACCAATCAGCTGATCTATGCAAC 119439

RESULT 11
AC084290/C
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

AC084290 55001 bp DNA linear PRI 28-MAR-2002
Homo sapiens 12 BAC RP11-568G5 (Rowe) Park Cancer Institute Human
BAC library complete sequence.
AC084290
AC084290.15 GI:19774277
HTG.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 55001)
Muzny, D.M., Adams, C., Adio-Oduola, B., All-ouman, F.R., Allen, C.,
Alsbrooks, S.L., Amaralunga, H.C., Are, J.R., Ayele, M., Banks, T.,
Barbieri, J., Benton, J., Bimaye, K., Blankenburg, K., Bonnin, D.,
Bouck, J., Bowles, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,
Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
Chen, G., Chen, R., Chen, Z., Chin, D., Chowdhry, I., Christopoulos, C.,
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
Davila, M.L., Davis, C., Davy-Carroll, L., Dedetich, D.A.,
Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, K.H.,
Doutlaw, K.J., Draper, H., Dugan-Rocha, S., Durkin, K.J.,
Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Emertling, S.,
Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P.,
Francz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N.,
Gill, R., Gorrell, J.H., Guereva, M., Gunaratne, P., Hale, S.,
Hamilton, K., Han, J., Harris, K., Hart, M., Havlak, P.,
Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hognes, M.,
Holloway, C., Hollins, B., Honsi, F., Howard, S., Huber, J., Hulyk, S.,
Hume, J., Ioshikhes, I., Jackson, L.E., Jacobson, B., Jia, Y.,
Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U.,
King, L., Kovach, J., Kovar, C., Kratoch, J., Kureishi, A., Landry, N.,
Leal, B., Lee, E., Lewis, L.C., Lewis, J., Li, J., Li, Z., Lichargos, O.,
Lien, C., Liu, J., Liu, W., Louie, H., Lozano, R., Lu, X.,
Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Magna, P.,
Marin, I., Martin, R., Martindale, A., Martinez, B., Massey, E.,
Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Metcalf, S.,
Metzker, M., Miller, A., Miner, G., Miner, Z., Mitchell, T.,
Mohabab, K., Montemery, K.T., Morgan, M., Morris, S., Moser, M.,
Neal, D., Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, B., Nockenwo, S., Ogun, M., Okunolu, G.,
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
Rivers, M., Rojas, A., Rojudo, I., Rolfe, M., Ruiz, S., Savery, G.,
Scheerer, S., Scott, G., Shen, H., Shim, C., Shoshitari, N., Sisson, I.,
Sodergren, E., Sonike, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Tabot, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tatney, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Uemini, K., Vaquez, L., Vera, V., Villalob, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watling, K.,
Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K.,
Wu, C., Wu, J., Wu, Y.F., Zhou, J., Zorilla, S., Kucherlapati, R.,
Weinstock, G. and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 55001)
REFERENCE
AUTHORS
TITLE
JOURNAL

Submitted (21-OCT-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 55001)
REFERENCE
AUTHORS
TITLE
JOURNAL

Submitted (27-MAR-2002) Human Genome Sequencing Center, Department

Of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 55001)
Morley, K.C.
Direct Submission
Submitted (28-VAR-2002) Human Genome Sequencing Center, Department

COMMENT On Mar 28, 2002 this sequence version replaced gi:19747063
 INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
 gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the features listing.

ANNOTATION OF FEATURES:
STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green unpublished.) for Human and Mouse sequences.

(Nuc. Acids Res. 25:3389-3402) similarly are identified by BLAST Genes and Region of sequence similarity are identified by BLAST EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT

FEATURES	Location/Qualifiers
source	1. .55001

```

misc_feature
1. .1301
/note="overlaps bases 177088. .178388 of clone AC079630"
/function="clone overlap"
35. .229
/rpc_family="MIR"
repeat_region
1002. .1023
/rpc_family="AT_rich"
repeat_region
1092. .1227
/rpc_family="MIR"
repeat_region
complement(2156. .2209)
/rpc_family="L2"
3563. .3650
/rpc_family="(CCGTA)n"
repeat_region
complement(3688. .3765)
/rpc_family="L2"
repeat_region
complement(3840. .3908)
/rpc_family="LIMD3"
repeat_region
complement(4234. .4464)
/rpc_family="MIR"

```

repeat_region	complement(4847, .4914) /rpt_family="MIR"
repeat_region	4938, .5053 /rpt_family="L2"
repeat_region	6350, .6660 /rpt_family="ALUSg" complement(7309, .7743) /rpt_family="LTR43"
repeat_region	7744, .9144 /rpt_family="L1P4" complement(9151, .9317) /rpt_family="LTR43"
repeat_region	10587, .10846 /rpt_family="L1M4" 10912, .11015 /rpt_family="L1P"
repeat_region	11098, .11271 /rpt_family="L1MB" 11450, .11635 /rpt_family="MLT11"
repeat_region	11941, .12179 /rpt_family="MIR"
repeat_region	12797, .12828 /rpt_family="CATATA"n complement(12975, .13087) /rpt_family="FLAM_C"
repeat_region	13724, .13737 /rpt_family="AT_rich"
repeat_region	14058, .14219 /rpt_family="MER45"
repeat_region	14411, .14545 /rpt_family="MIR" complement(15058, .15127) /rpt_family="MIR"
repeat_region	16186, .16413 /rpt_family="MIR"
repeat_region	complement(17624, .17922) /rpt_family="ALUub"
repeat_region	complement(18309, .18606) /rpt_family="ALUSx"
repeat_region	18844, .18854 /rpt_family="AT_rich"
repeat_region	complement(18855, .19135) /rpt_family="ALUub"
repeat_region	19136, .19158 /rpt_family="AT_rich"
repeat_region	complement(20660, .20753) /rpt_family="MER91C"
repeat_region	22179, .22680 /rpt_family="(TAAA)n"
repeat_region	22180, .22358 /rpt_family="ALUJo"
repeat_region	22359, .22632 /rpt_family="ALUub"
repeat_region	22633, .22680 /rpt_family="(TAAA)n"
repeat_region	22682, .22789 /rpt_family="ALUJo"
repeat_region	22796, .23099 /rpt_family="ALUy"
repeat_region	complement(23147, .23327) /rpt_family="MIR"

Query Match	60.0%	Score 24;	DB 9;	Length 55001;
Best Local Similarity	84.4%	Pred. No. 48;		
Matches 27; Conservative	0;	Mismatches 5;	Indels 0;	Gaps 0;

Qy 4 ATACCTACTCCCAACAATTCATTACTCAT 35
 | | | | | | | | | | | | | | |
Db 47598 AACTTCTACTCCCAACAATTCATTGCTACTGT 47667

RESULT 12

DEFINITION	Accession	Version	Keywords	Source	Organism
Papio anubis clone RP41-375M17, WORKING DRAFT SEQUENCE, 3 ordered pieces.	AC148514	2	GI:45680463	HTGS PHASE2; HTGS DRAFT.	Papio anubis (olive baboon)
REFERENCE	1 (bases 1 to 171940)				
AUTHORS	Antoniellis, A., Ayele, K., Benjamin, B., Blakeley, R. W., Coleman, H., Dahl, N., Engle, J., Granite, S., Guan, X., Gupta, J., Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Hu, P., Hurle, B., Idol, J. R., Jones, C., Karlins, E., Kim, H., Kwong, P., Latic, P., Larson, S., Lee-Lin, S.-Q., Legaspi, R., Maduro, Q. L., Maduro, V. B., Margulies, E. H., Masiello, C., Maskeri, B., McQuinn, J., Mullikin, J. C., Pasquitan, C., Portnoy, M. E., Prasad, A., Puri, O., Reddy-Dugue, N., Schandier, K., Schuler, M. G., Shah, K., Sison, C., Stantrop, S., Thomas, J. W., Thomas, P. J., Tsipouri, V., Vogt, J. L., Wehrby, K. D., Young, A. and Green, E. D.				
TITLE	NISC Comparative Sequencing Initiative				
REFERENCE	2 (bases 1 to 171940)				
AUTHORS	Green, E. D.				
TITLE	Direct Submision				
JOURNAL	Submitted (10-MAR-2004) NIH Intramural Sequencing Center, 8717 Grovemont Circle, Gaithersburg, MD 20877, USA				
REFERENCE	3 (bases 1 to 171940)				
AUTHORS	Green, E. D.				
TITLE	Direct Submision				
JOURNAL	Submitted (24-MAR-2004) NIH Intramural Sequencing Center, 8717 Grovemont Circle, Gaithersburg, MD 20877, USA				
COMMENT	On Mar 24, 2004 this sequence version replaced gi:45332443.				
	----- Genome Center				
	Center: NIH Intramural Sequencing Center				
	Center code: NISC				
	Web site: http://www.nisc.nih.gov				
	Contact: nisc.zoo@hgrl.nih.gov				
	----- Project Information				
	Center project name: fhm				
	Center clone name: 375M17				

	The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.				
	----- Summary Statistics				
	Sequencing vector: plasmid, n/a, 100% of reads				
	Chemistry: Dye-terminator Big Dye; 100% of reads				
	Assembly program: Phrap; version 0.990319				
	Consensus quality: 171395 bases at least Q40				
	Consensus quality: 171621 bases at least Q30				
	Consensus quality: 171701 bases at least Q20				
	Insert size: 173000; agarose-fp				
	Insert size: 171740; sum-of-contigs				
	Quality coverage: 8.27x in Q20 bases; agarose-fp				
	Quality coverage: 8.33x in Q20 bases; sum-of-contigs				

	* NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces are believed to be correct as given, however the sizes of the gaps between them are based on estimates that have				

```

* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
* 1 10148: contig of 10148 bp in length
* 10149 10248: gap of unknown length
* 10249 103023: contig of 92775 bp in length
* 103024 103123: gap of unknown length
* 103124 171940: contig of 68617 bp in length.
FEATURES
    source
        1..171940
            /organism="Papio anubis"
            /mol_type="genomic DNA"
            /db_xref="taxon:9555"
            /clone="RP41-375M17"
            /clone_lib="RP41"
            /note="BAC resource: http://bacpac.chori.org/"
        1..10148
            /note="assembly_fragment"
            clone_end:SP6
            vector_side:left"
        10249..103023
            /note="assembly_fragment"
            103124..171940
            /note="assembly_fragment"
            clone_end:T7
            vector_side:right"
ORIGIN
Query Match      60.0%; Score 24; DB 2; Length 171940;
Best Local Similarity 84.4%; Pred. No. 42;
Matches 27; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY      4 ATACCTACCTCCCAACATCATCTACTCAT 35
DB      33556 AACTTACTCTCCCAACATCATCTACTCAT 33525

RESULT 14
AC148501/c 193735 bp DNA linear HTG 09-JUL-2004
LOCUS      AC148501
DEFINITION Callitrix jacchus clone CH259-368E20, WORKING DRAFT SEQUENCE, 11
ordered pieces.
ACCESSION  AC148501
VERSION     AC148501.3 GI:50080339
KEYWORDS   HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE     Callitrix jacchus (white-tufted-ear marmoset)
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae;
Callitrix.
1 (bases 1 to 193735)
Antoneillis,A., Aylele,K., Benjamin,B., Blakesley,R.W.,
Boutifard,G.G., Brinkley,C., Brooks,S., Chu,G., Coleman,B.,
Coleman,H., Dakl,N., Engle,J., Guan,X., Gupta,J., Haghighi,P.,
Han,J., Hansen,N., Ho,S.-L., Hu,P., Hurle,B., Idol,J.R., Jones,C.,
Karling,E., Kim,H., Kwong,P., Laric,P., Larson,S., Lee-Ilh,S.-O.,
Legasse,J.R., Medden,M., Maduro,Q.L., Maduro,V.B., Margulies,E.H.,
Masseillo,C., Maskell,B., Mcowell,J., Mullikin,J.C., Pagnutigan,C.,
Park,M., Portnoy,M.E., Prasad,A., Puri,O., Reddi-Dugue,N.,
Schandler,K., Schueler,M.G., Shah,K., Sison,C., Stantirpoo,S.,
Thomas,J.W., Thomas,P.J., Teipouri,V., Vogt,J.L., Wetherby,K.D.,
Young,A. and Green,E.D.
MISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 193735)
REFERENCE
JOURNAL    Submitted (10-MAR-2004) NIH Intramural Sequencing Center, 8717
Grovenom Circle, Gaithersburg, MD 20877, USA
AUTHORS    Green,E.D.
REFERENCE  3 (bases 1 to 193735)
AUTHORS    Direct Submission
TITLE      Direct Submission
JOURNAL    Submitted (10-MAR-2004) NIH Intramural Sequencing Center, 8717
Grovenom Circle, Gaithersburg, MD 20877, USA
REFERENCE  3 (bases 1 to 193735)
AUTHORS    Direct Submission
TITLE      Direct Submission

```

```

JOURNAL    Submitted (09-JUL-2004) NIH Intramural Sequencing Center, 8717
Grovenom Circle, Gaithersburg, MD 20877, USA
On Jul 9, 2004 this sequence version replaced gi:46430810.
COMMENT
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoemhri.nih.gov
----- Project Information
Center project name: ITW
Center clone name: 368E20

The sequence data in this record represents an 'enhanced'
version of a Phase 2 submission. Specifically, the indicated
order and orientation of each sequence contig has been
established using one or more of the following: read-pair
data from individual subclones, overlaps with neighboring
clones, alignment with available reference sequence (e.g.,
human), and/or confirmation by PCR testing. In addition,
the sequence assembly is based on at least 8X average
coverage in Q20 bases and has been reviewed to rule out
gross misassemblies, the low-quality ends of sequence
contigs have been trimmed away, and each base is associated
with a Phrap-derived quality score.
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 190717 bases at least Q40
Consensus quality: 191709 bases at least Q30
Consensus quality: 192323 bases at least Q20
Insert size: 14400; agarose-fp
Insert size: 192735; sum-of-contigs
Quality coverage: 13.33x in Q20 bases; sum-of-contigs
Quality coverage: 9.96x in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 11 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
* 1 29309: contig of 29309 bp in length
* 29310 29409: gap of unknown length
* 29410 37712: contig of 8303 bp in length
* 37713 37812: gap of unknown length
* 37813 40472: contig of 2660 bp in length
* 40473 40572: gap of unknown length
* 40573 45724: contig of 5152 bp in length
* 45725 45824: gap of unknown length
* 45825 47844: contig of 2020 bp in length
* 47845 47944: gap of unknown length
* 47945 61027: contig of 13083 bp in length
* 61028 61127: gap of unknown length
* 61128 63312: contig of 2085 bp in length
* 63313 63312: gap of unknown length
* 63313 70560: contig of 7248 bp in length
* 70561 70660: gap of unknown length
* 70661 73658: contig of 2998 bp in length
* 73659 73759: gap of unknown length
* 73759 103653: contig of 29895 bp in length
* 103654 103753: gap of unknown length
* 103754 193735: contig of 89982 bp in length.
FEATURES
    source
        1..193735
            /organism="Callitrix jacchus"
            /mol_type="genomic DNA"
            /db_xref="taxon:9483"
            /clone="CH259-368E20"
            /clone_lib="CH259"

```

misc_feature /note="BAC resource: <http://bacpac.chori.org/>"
1..29309
/note="assembly_fragment"
clone_end:T7
vector_side:left"
29410..37712
/note="assembly_fragment"
37813..40472
/note="assembly_fragment"
40573..45724
/note="assembly_fragment"
45825..47844
/note="assembly_fragment"
47945..61027
/note="assembly_fragment"
61128..63212
/note="assembly_fragment"
63313..70560
/note="assembly_fragment"
70661..73658
/note="assembly_fragment"
73759..103653
/note="assembly_fragment"
103754..193735
/note="assembly_fragment"
clone_end:SP6
vector_side:right"
127712..193735
/note="clone overlaps with GenBank Accession Number
AC148551 clone CH259-256024 (center project name fvv)"

ORIGIN

Query Match 59.0%; Score 23.6; DB 2; Length 193735;
Best Local Similarity 76.3%; Pred. No. 60;
Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2 GGATACCTACTCCGAACATCCATCTTACTCATGCAA 39

DB 89109 GGAAATCACTCCAAAATAATCCCTCAAAATCATGCAA 89072

RESULT 15
AC008897 117026 bp DNA linear PRI 07-NOV-2000
LOCUS AC008897 Homo sapiens chromosome 5 clone CTD-2235C13, complete sequence.
AC008897
AC008897.7 GI:11119448
HTG.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 117026)
DOE Joint Genome Institute and Stanford Human Genome Center.
Direct Submision
Unpublished
2 (bases 1 to 117026)
DOE Joint Genome Institute.
Direct Submision
Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 117026)
DOE Joint Genome Institute and Stanford Human Genome Center.
Direct Submision
Submitted (07-NOV-2000) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
On Nov 7, 2000 this sequence version replaced gi:7711370.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www-shgc.stanford.edu
Quality: Phrap Quality >=40 99.5% of Sequence;
Estimated Total Number of Errors is 0.5.

STS Content:
WI-9815 G05428
SHGC-16751 G15410
WI-7171 G06431.
Location/Qualifiers
1..117026
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTD-2235C13"

ORIGIN

Query Match 57.5%; Score 23; DB 9; Length 117026;
Best Local Similarity 74.4%; Pred. No. 1.1e+02;
Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCGAACATCCATCTTACTCATGCAA 39

DB 72634 AGGAACGGATTAACTAATCCATCTTCTCAAGCAA 72672

Search completed: September 16, 2005, 03:01:40
Job time : 136.222 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using bw model

Run on: September 15, 2005, 08:10:17 ; Search time 19.3656 Seconds
(without alignments)
12227.321 Million cell updates/sec

Title: US-10-713-381-2_COPY_1239_1278

Perfect score: 40

Sequence: 1 aggatacctaccctcccaacaacatccatcctaccgagcaac 40

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: genebegn1980s:*
2: genebegn1990s:*
3: genebegn2000s:*
4: genebegn2001as:*
5: genebegn2001bs:*
6: genebegn2002as:*
7: genebegn2002bs:*
8: genebegn2003as:*
9: genebegn2003bs:*
10: genebegn2003cs:*
11: genebegn2003ds:*
12: genebegn2004as:*
13: genebegn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	40	100.0	40	5	AAH76337	Aah76337 Z. mays M
2	40	100.0	50	5	AAH76336	Aah76336 Z. mays M
3	40	100.0	158	5	AAH76334	Aah76334 Z. mays M
4	40	100.0	255	5	AAH76340	Aah76340 Z. mays M
5	40	100.0	1394	2	AAH7408	Aax7408 Zea mays
6	40	100.0	1394	2	AAH7409	Aax7409 Zea mays
7	40	100.0	1394	5	AAH76332	Aah76332 Z. mays M
8	40	100.0	1394	5	AAH76333	Aah76333 Z. mays M
9	22.4	56.0	226215	11	ACN45146	Acn45146 Human gen
10	21.6	54.0	4110	8	ABZ10202	Abz10202 Haematopo
11	21.6	54.0	6309	6	ABL23204	Ab132304 Human imm
12	21.6	54.0	7110	10	ADB54282	Add54282 Pretreat
13	21.6	54.0	7110	10	ADB84196	Ad84196 Human lyl
14	21.6	54.0	7110	13	ADS89580	Ads89580 Oligonuc
15	21.4	53.5	1341	6	ABZ13750	Abz13750 Arabidops
16	21.4	53.5	6681	6	ABL2155	Ab12155 Human imm
17	21.4	53.5	6681	6	ABL54304	Ab154304 Chemicall
18	21.4	53.5	8166	6	ABL132928	Ab132928 Human imm
19	21.4	53.5	8166	6	AA863328	Aa863328 Chemicall
20	21.4	53.5	24401	4	ABLO3396	Ablo3396 Drosophil

21	21	52.5	2000	12	ADJ41594	Adj41594 Plant con
22	21	52.5	2792	12	ADM98941	Adm98941 Diterpene
23	21	52.5	3117	12	ADM98942	Adm98942 Diterpene
24	21	52.5	5507	8	ABZ10204	Abz10204 Haematopo
25	21	52.5	5507	8	ABZ10058	Abz10058 Haematopo
26	21	52.5	8404	4	AA646500	Aa646500 Tumour su
27	21	52.5	8404	6	ABL33595	Ab133595 Human imm
28	21	52.5	8404	10	ADB54216	Adb54216 Pretreate
29	21	52.5	8404	10	ADB54088	Adb54088 Pretreate
30	21	52.5	8404	10	ADB84076	Adb84076 Human lym
31	21	52.5	8404	10	ADB84152	Adb84152 Human lym
32	21	52.5	8404	13	ADS89242	Ads89242 Oligonuc
33	21	52.5	8404	13	ADS89516	Ads89516 Oligonuc
34	21	52.5	10286	4	AA645308	Aa645308 Chemical
35	21	52.5	10286	6	ABK28147	Abk28147 DNA trans
36	21	52.5	63115	10	ADC85419	Adc85419 Mouse tle
37	21	52.5	63294	9	ADA02939	Ada02939 Mouse tle
38	21	52.5	63294	10	ADB72677	Adb72677 Mouse tle
39	21	52.5	63294	12	ADM74534	Adm74534 Murine ca
40	21	52.5	106315	11	ACN43966	Acn43966 Human gen
41	20.8	891	8	ABZ51893	Abz51893 Aspergill	
42	20.8	52.0	6620	4	AA645488	Aa645488 Chemical
43	20.8	52.0	6620	6	ABK28416	Abk28416 DNA trans
44	20.8	52.0	6620	6	ABN80279	Abn80279 Human che
45	20.8	52.0	110000	5	AA161373_1	Continuation (2 of

ALIGNMENTS

RESULT 1	AAH76337	standard; DNA; 40 BP.
ID	AAH76337	standard; DNA; 40 BP.
AC	AAH76337	
XX	AAH76337	
DT	29-OCT-2001	(first entry)
XX	29-OCT-2001	(first entry)
DE	Z. mays Ms45 male tissue-preferred regulatory region fragment.	
XX	Z. mays Ms45 male tissue-preferred regulatory region fragment.	
KM	Ms45; male tissue; regulatory region; transcription; male fertility;	
XX	hybrid seed; ds.	
OS	Zea mays.	
XX	Zea mays.	
PN	WQ200160997-A2.	
XX	WQ200160997-A2.	
PD	23-AUG-2001.	
XX	23-AUG-2001.	
PF	13-FEB-2001; 2001WO-US004527.	
XX	13-FEB-2001; 2001WO-US004527.	
PR	15-FEB-2000; 2000US-00504487.	
XX	15-FEB-2000; 2000US-00504487.	
PA	(PION-) PIONEER HI-BRED INT INC.	
XX	(PION-) PIONEER HI-BRED INT INC.	
PI	Albertsen MC, Fox TW, Garnaat CW, Hufman G, Kendall TL;	
XX	Albertsen MC, Fox TW, Garnaat CW, Hufman G, Kendall TL;	
DR	WPI; 2001-514772/56.	
XX	WPI; 2001-514772/56.	
PT	A male tissue-preferred regulatory region comprising nucleotide sequences	
XX	essential for initiating transcription of the Ms45 gene useful for	
PT	mediating fertility in a male plant.	
XX	mediating fertility in a male plant.	
PS	Claim 14; Page 32; 50p; English.	
XX	Claim 14; Page 32; 50p; English.	
CC	The invention provides a male tissue-preferred regulatory region (I)	
XX	comprising nucleotide sequences essential for initiating transcription of	
CC	the Ms45 gene. A method of mediating male fertility in a plant is	
XX	provided that involves introducing an expression vector comprising a	
CC	promoter operably linked to (I) into a plant where the exogenous gene	
XX	impacts male fertility of the plant and (I) controls expression of the	
CC	exogenous gene. A method of producing hybrid seeds is also provided. The	
XX	present sequence represents a DNA fragment upstream of the TATA box of a	

```
CC 2. mays Ms45 male-tissue preferred regulatory region nucleotide sequence
XX Sequence 40 BP; 14 A; 14 C; 3 G; 9 T; 0 U; 0 Other;
SQ

Query Match      100.0%; Score 40; DB 5; Length 40;
Best Local Similarity 100.0%; Pred. No. 3.6e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AGGATACCTACTCCCAACATCCATCTTACTCATGCAAC 40
   |||||||
DB 1 AGGATACCTACTCCCAACATCCATCTTACTCATGCAAC 40

RESULT 2
AAH76336
ID AAH76336 standard; DNA; 50 BP.
XX
AC AAH76336;
XX
DT 29-OCT-2001 (first entry)
XX
DE 2. mays Ms45 male tissue-preferred regulatory region fragment.
XX
KW Ms45; male tissue; regulatory region; transcription; male fertility;
KW hybrid seed; ds.
XX
OS Zea mays.
XX
PN WO200160997-A2.
XX
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
DR WPI; 2001-514772/56.
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the Ms45 gene useful for
PT mediating fertility in a male plant.
XX
PS Claim 5; Page 47; 50pp; English.
XX
CC The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the Ms45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a DNA fragment -72 to -111 bases upstream of
CC the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
CC nucleotide sequence
XX
SQ Sequence 50 BP; 18 A; 17 C; 5 G; 10 T; 0 U; 0 Other;

Query Match      100.0%; Score 40; DB 5; Length 50;
Best Local Similarity 100.0%; Pred. No. 3.7e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AGGATACCTACTCCCAACATCCATCTTACTCATGCAAC 40
   |||||||
DB 1 AGGATACCTACTCCCAACATCCATCTTACTCATGCAAC 50

RESULT 3
AAH76334
ID AAH76334 standard; DNA; 158 BP.
```

```
XX AAH76334;
AC
XX 29-OCT-2001 (first entry)
DT
XX 2. mays Ms45 male tissue-preferred regulatory region fragment.
DE
XX Ms45; male tissue; regulatory region; transcription; male fertility;
KW hybrid seed; ds.
XX
OS Zea mays.
XX
PN WO200160997-A2.
XX
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
DR WPI; 2001-514772/56.
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the Ms45 gene useful for
PT mediating fertility in a male plant.
XX
PS Claim 5; Page 47; 50pp; English.
XX
CC The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the Ms45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a DNA fragment -38 to -195 bases upstream of
CC the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
CC nucleotide sequence
XX
SQ Sequence 158 BP; 41 A; 50 C; 21 G; 46 T; 0 U; 0 Other;

Query Match      100.0%; Score 40; DB 5; Length 158;
Best Local Similarity 100.0%; Pred. No. 4.4e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AGGATACCTACTCCCAACATCCATCTTACTCATGCAAC 40
   |||||||
DB 86 AGGATACCTACTCCCAACATCCATCTTACTCATGCAAC 125

RESULT 4
AAH76340
ID AAH76340 standard; DNA; 255 BP.
XX
AC AAH76340;
XX
DT 29-OCT-2001 (first entry)
XX
DE 2. mays Ms45 promoter fragment.
XX
KW Ms45; male tissue; regulatory region; transcription; male fertility;
KW hybrid seed; promoter; ds.
XX
OS Zea mays.
XX
PN WO200160997-A2.
XX
PD 23-AUG-2001.
```


PF 13-FEB-2001; 2001MO-US004527.
XX
XX 15-FEB-2000; 2000US-00504487.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX WPI; 2001-514772/56.
XX
XX A male tissue-preferred regulatory region comprising nucleotide sequences
XX essential for initiating transcription of the Ms45 gene useful for
XX mediating fertility in a male plant.
XX
XX Example 5; Fig 8; 50pp; English.
XX
XX The invention provides a male tissue-preferred regulatory region (I)
XX comprising nucleotide sequences essential for initiating transcription of
XX the Ms45 gene. A method of mediating male fertility in a plant is
XX provided that involves introducing an expression vector comprising a
XX promoter operably linked to (I) into a plant where the exogenous gene
XX impacts male fertility of the plant and (I) controls expression of the
XX exogenous gene. A method of producing hybrid seeds is also provided. The
XX present sequence represents a Z. mays Ms45 promoter fragment
XX
SQ Sequence 255 BP; 59 A; 86 C; 39 G; 71 T; 0 U; 0 Other;

Query Match 100.0%; Score 40; DB 5; Length 255;
Best Local Similarity 100.0%; Pred. No. 4.7e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCTCCAAACATTCATCTTACTCATGCAAC 40
DB 99 AGGATACCTACTCTCCAAACATTCATCTTACTCATGCAAC 138

RESULT 5
AA07408
ID AAX07408 standard; DNA; 1394 BP.
XX
XX AAX07408;
XX
XX 08-JUN-1999 (first entry)
XX
XX Zea mays Ms45 male tissue-preferred regulatory region.
XX
XX Ms45; male; tissue-preferred; regulatory region; plant cells;
XX plant tissue; differentiated; maize; hybrid seed; fertility; ss.
XX
XX Zea mays.
XX
XX WO9859061-A1.
XX
XX 30-DEC-1998.
XX
XX 19-JUN-1998; 98WO-US012895.
XX
XX 23-JUN-1997; 97US-00880499.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Albertsen MC, Fox TW, Garnaat CW, Huffman GA, Kendall TL;
XX WPI; 1999-105628/09.
XX
XX New nucleic acid encoding a Ms45 male tissue-preferred regulatory region
XX - useful in mediating plant fertility, especially hybrid seed production.
XX
XX Claim 2; Page 22-23; 39pp; English.
XX
XX The sequence is that encoding an Ms45 male tissue-preferred regulatory
XX region. It may be used in the construction of a vector for a method of
XX producing exogenous genes in a male tissue-preferred manner, which is

CC useful in restoring or conferring fertility, such as in hybrid seed
CC production. In conferring fertility, a monocot/dicot plant is transformed
CC with the exogenous nucleotide sequence (a male sterility gene, preferably
CC Ms45), which encodes a product selected from auxins, rolB and diptheria
CC toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
CC and infertile plants
XX
SQ Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;

Query Match 100.0%; Score 40; DB 2; Length 1394;
Best Local Similarity 100.0%; Pred. No. 6.1e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCTCCAAACATTCATCTTACTCATGCAAC 40
DB 1239 AGGATACCTACTCTCCAAACATTCATCTTACTCATGCAAC 1278

RESULT 6
AA07409
ID AAX07409 standard; DNA; 1394 BP.
XX
XX AAX07409;
XX
XX 08-JUN-1999 (first entry)
XX
XX Zea mays Ms45 male tissue-preferred regulatory region.
XX
XX Ms45; male; tissue-preferred; regulatory region; plant cells;
XX plant tissue; differentiated; hybrid seed; fertility; ss.
XX
XX Zea mays.
XX
XX WO9859061-A1.
XX
XX 30-DEC-1998.
XX
XX 19-JUN-1998; 98WO-US012895.
XX
XX 23-JUN-1997; 97US-00880499.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Albertsen MC, Fox TW, Garnaat CW, Huffman GA, Kendall TL;
XX WPI; 1999-105628/09.
XX
XX New nucleic acid encoding a Ms45 male tissue-preferred regulatory region
XX - useful in mediating plant fertility, especially hybrid seed production.
XX
XX Claim 3; Page 23-24; 39pp; English.
XX
XX The sequence is that encoding an Ms45 male tissue-preferred regulatory
XX region. It may be used in the construction of a vector for a method of
XX producing exogenous genes in a male tissue-preferred manner, which is
XX useful in restoring or conferring fertility, such as in hybrid seed
XX production. In conferring fertility, a monocot/dicot plant is transformed
XX with the exogenous nucleotide sequence (a male sterility gene, preferably
XX Ms45), which encodes a product selected from auxins, rolB and diptheria
XX toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
XX and infertile plants
XX
SQ Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;

Query Match 100.0%; Score 40; DB 2; Length 1394;
Best Local Similarity 100.0%; Pred. No. 6.1e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCTCCAAACATTCATCTTACTCATGCAAC 40
DB 1239 AGGATACCTACTCTCCAAACATTCATCTTACTCATGCAAC 1278

```
RESULT 7
AAH76332
ID AAH76332 standard; DNA; 1394 BP.
XX
AC AAH76332;
XX
DT 29-OCT-2001 (first entry)
XX
DE Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.
XX
KM Ms45; male tissue; regulatory region; transcription; male fertility;
KW hybrid seed; ds.
XX
OS Zea mays.
XX
PN WO200160997-A2.
XX
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
DR WPI; 2001-514772/56.
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the Ms45 gene useful for
PT mediating fertility in a male plant.
XX
PS Claim 4; Page 46; 50pp; English.
XX
CC The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the Ms45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a nucleic acid sequence encoding an Ms45 male
CC -tissue preferred regulatory region from Z. mays
XX
SQ Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 40; DB 5; Length 1394;
Best Local Similarity 100.0%; Pred. No. 6.1e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 AGGATACCTACTCCCAACAATCCATCTTACTGATGCAAC 40
DB 1239 AGGATACCTACTCCCAACAATCCATCTTACTGATGCAAC 1278
XX
RESULT 8
AAH76333
ID AAH76333 standard; DNA; 1394 BP.
XX
AC AAH76333;
XX
DT 29-OCT-2001 (first entry)
XX
DE Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.
XX
KM Ms45; male tissue; regulatory region; transcription; male fertility;
KW hybrid seed; ds.
XX
OS Zea mays.
XX
PN WO200160997-A2.
XX
```

```
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
DR WPI; 2001-514772/56.
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the Ms45 gene useful for
PT mediating fertility in a male plant.
XX
PS Claim 4; Page 47; 50pp; English.
XX
CC The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the Ms45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a nucleic acid sequence encoding an Ms45 male
CC -tissue preferred regulatory region from Z. mays
XX
SQ Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 40; DB 5; Length 1394;
Best Local Similarity 100.0%; Pred. No. 6.1e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 AGGATACCTACTCCCAACAATCCATCTTACTGATGCAAC 40
DB 1239 AGGATACCTACTCCCAACAATCCATCTTACTGATGCAAC 1278
XX
RESULT 9
ACN45146
ID ACN45146 standard; DNA; 226215 BP.
XX
AC ACN45146;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human genomic sequence hCG1639824.
XX
KW Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003073826-A2.
XX
PD 12-SEP-2003.
XX
PF 28-FEB-2003; 2003WO-US006235.
XX
PR 01-MAR-2002; 2002US-00087192.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW;
XX
DR WPI; 2003-328604/31.
XX
PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
PT comprises a nucleotide sequence.
XX
PS Claim 1; SEQ ID NO 1948; Opp; English.
XX
CC The present invention relates to novel DNA and protein sequences which
```

CC are associated with carcinomas. The sequences are useful for: (i) for
CC screening drug candidates; (ii) for screening of bioactive agent capable
CC of binding to CarcinoMa Associated Protein (CAP); (iii) for screening of
CC a bioactive agent capable of modulating the activity of CAP; (iv) for
CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a bioprobe;
CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
CC determining CarcinoMa Associated (CA) gene copy number. In addition, these
CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
CC carcinoma including lymphoma. The present sequence is one such CA coding
CC sequence. Note: This patent is an equivalent to basic patent
US2002182586A1, for which no sequence data was published

SQ Sequence 226215 BP; 65331 A; 42804 C; 45218 G; 72862 T; 0 U; 0 Other;

Query Match	56.0%	Score 22.4;	DB 11;	Length 226215;
Best Local Similarity	72.5%;	Pred. No. 1.2e+02;		
Matches 29; Conservative	0;	Mismatches 11;	Indels 0;	Gaps 0;

```

Oy      1 AGGATACCTTACTCTCCAAACAATCCATCTTACTCATGCAAC 40
          |||||  |||||  |||||  |||||  |||||  |||||
Db      183206 AGAATTCTGCTTTCCAACAACAGCATCTTACTCTTTTAC 183245

```

RESULT 10
ABZ10202/c
ID ABZ10202 standard; DNA; 4110 BP.

AC ABZ10202;

DT 16-JAN-2003 (first entry)

DE Haematopoietic cell proliferation disorder related DNA sequence #342.

Human; haematopoietic cell proliferation disorder; cytostatic; gene therapy; lymphocytic leukemia; acute myelogenous leukaemia; cytosine methylation state; gene; ds.

Homo sapiens.

PN WO200277272-A2.

PD 03-OCT-2002.

PF 26-MAR-2002; 2002WO-EP003401.

PR 26-MAR-2001; 2001US-0278333P.

PA (EPIC-) EPIGENOMICS AG.

PI Berlin K, Braun A, Distler J, Guetzig D, Howe A, Mueller J;
PI Olek A, Pienhopfer C, Adorian P, Grabs G, Lesche R, Liu E;
PI Lewin A, Lipschier E, Maier S, Model F, Mueller V, Otto T, Peter C;
PI Schwope I, Ziebach H;

DR WPI; 2003-018942/01.

PT Detecting and differentiating between hematopoietic cell proliferative
PT disorders, comprises contacting a target nucleic acid with a reagent that
PT distinguishes between methylated and non-methylated CpG dinucleotides.
XX
PS Claim 28; SEQ ID NO 342; 11pp; English.

CC The present invention describes a method

CC The present invention describes a method for detecting and
CC differentiating between haematopoietic cell proliferative disorders
CC associated with at least 1 gene and/or their regulatory regions in a
CC subject. The method comprises contacting a target nucleic acid in a
CC biological sample obtained from the subject with at least 1 reagent,
CC which distinguishes between methylated and non-methylated CpG
CC dinucleotides within the target nucleic acid. AB209861 to AB21118
CC represent specifically claimed nucleotide sequences for the present
CC invention. Oligonucleotides from the present invention can be used: for

differentiating between healthy haematopoietic cells and proliferative disorder haematopoietic cells; for differentiating between acute lymphocytic leukaemia and acute myelogenous leukaemia; as probes for determining the cytosine methylation state and/or single nucleotide polymorphisms (SNPs) of haematopoietic cell proliferation disorder related sequences and their complements; and as primers for the amplification of haematopoietic cell proliferation disorder related sequences. The nucleotide sequences from the present invention can also be used for detecting a predisposition to, differentiation between subclasses, diagnosis, prognosis, treatment and/or monitoring of haematopoietic cell proliferative disorders. The present method enables highly specific classification of haematopoietic cell proliferative disorders allowing for improved and informed treatment of patients

SQ Sequence 4110 BP; 945 A; 0 C; 1132 G; 2033 T; 0 U; 0 Other;

Query Match	54.0%	Score 21.6;	DB 8;	Length 4110;
Best Local Similarity	85.7%;	Pred. No. 1.4e+02;		
Matches 24; Conservative	0;	Mismatches 4;	Indels 0;	Gaps 0;

QY **4** ATACCTACTCCCAACAATCCATCTTAC 31
| | | | | | | | | | | | | | |
Db 3083 AAACCTAACCCCAACAATCATCTTAC 3056

RESULT 11
ABL32304/c
ID ABL32304 standard; DNA; 6309 BP.

AC ABL32304;

DT 26-MAR-2002 (first entry)

DE Human immune system associated gene SEQ ID NO: 277.

KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiatherosclerotic; antihaemic; cytosatic; nootropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;

antiinflammatory; cancer; eye disease; arteriosclerosis; anas

KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease

3
2
1

XX

XX
173

XX	6	7	8	9
----	---	---	---	---

PF 02-JUL-2001; 2001WO-EP007537.

PR 30-JUN-2000; 2000DE-01032529.

XX	4	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	4
----	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	---

XX

PI Olek A, Piepenbrock C, Berlin K,

DR WPI; 2002-130909/17.

PT Nucleic acid comprising fragment of chemically modified gene, useful for
PT diagnosis and treatment of diseases associated with abnormal cytosine
PT methylation.

PS Claim 1; SEQ ID NO 277; 32pp + Sequence Listing; German

CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC

CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention
 XX
 SQ Sequence 6309 BP; 1758 A; 69 C; 1364 G; 3118 T; 0 U; 0 Other;

Query Match 54.0%; Score 21.6; DB 6; Length 6309;
 Best Local Similarity 75.0%; Pred. No. 1.5e+02;
 Matches 27; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 4 ATACCTACTCCCAACCAATCCATCTTACATGCAA 39
 DB 3335 ATACCTATCTCCCAATGACCAATCTCAATATCCAA 3300

RESULT 12
 ADB54282/C
 ID ADB54282 standard; DNA; 7110 BP.

AC ADB54282;

DT 04-DEC-2003 (first entry)

DE Pretreated genomic DNA region 206.

KW colon cell proliferative disorder; non methylated CpG dinucleotide;
 KW cytostatic; cancer; adenoma; carcinoma; cytosine methylation state; ds.

OS Unidentified.

XX WO2003072821-A2.

PD 04-SEP-2003.

PF 27-FEB-2003; 2003WO-EP002035.

PR 27-FEB-2002; 2002EP-00004551.

XX (EPIG-) EPIGENOMICS AG.

PA Adorjan P, Burger M, Maier S, Nimnrich I, Becker E, Lesche R;

PI Rujan T, Schmitt A;

XX WPI; 2003-731620/69.

PT Detecting and differentiating between colon cell proliferative disorders
 PT associated with a gene or its regulatory regions comprising contacting a
 PT target nucleic acid in a biological sample obtained from the subject with
 PT a reagent.

PS Claim 32; SEQ ID NO 338; 74pp; English.

CC The invention relates to a novel method for detecting and differentiating
 CC between colon cell proliferative disorders associated with at least one
 CC gene or its regulatory regions. The method comprises contacting a target
 CC nucleic acid in a biological sample obtained from the subject with at
 CC least one reagent or a series of reagents, where the reagent or series of
 CC reagents, distinguishes between methylated and non methylated CpG
 CC dinucleotides within the target nucleic acid. The molecules of the
 CC invention demonstrate cytostatic activity whilst the method may useful
 CC for detecting and differentiating between colon cell proliferative
 CC disorders, including cancers such as colon adenoma and colon carcinoma.
 CC The pNA (peptide nucleic acid)-oligomers are useful as probes for
 CC determining cytosine methylation state or single nucleotide
 CC polymorphisms. The current sequence is that of the pretreated genomic DNA
 CC region of the invention. This sequence is not shown within the
 CC specification but is taken from Wipoweb.

XX Sequence 7110 BP; 1818 A; 0 C; 1721 G; 3571 T; 0 U; 0 Other;

Query Match 54.0%; Score 21.6; DB 10; Length 7110;
 Best Local Similarity 85.7%; Pred. No. 1.5e+02;
 Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 ATACCTACTCCCAACCAATCCATCTTAC 31
 DB 3083 AAACCTAACCCCAACCAATCCATCTTAC 3056

RESULT 13
 ADE84196/C
 ID ADE84196 standard; DNA; 7110 BP.

AC ADE84196;

DT 29-JAN-2004 (first entry)

DE Human lymphoid cell proliferative disorder gene derived DNA #132.

KW ds; lymphoid cell proliferative disorder; methylation;
 KW methylated CpG dinucleotide; single nucleotide polymorphism; SNP;
 KW diffuse large B-cell lymphoma; mantle cell lymphoma;
 KW chronic lymphocytic leukemia; small lymphocytic lymphoma;
 KW follicular lymphoma; diagnosis; prognosis.

OS Homo sapiens.

XX WO2003044226-A2.

PD 30-MAY-2003.

PF 25-NOV-2002; 2002MO-EP013265.

PR 23-NOV-2001; 2001DE-01057491.

PR 28-DEC-2001; 2001DE-01064501.

XX (EPIG-) EPIGENOMICS AG.

PI Burger M, Caldwell C, Genc B, Becker E, Maier S, Nimnrich I;

XX WPI; 2003-457621/43.

PT Detecting and differentiating between lymphoid cell proliferative
 PT disorders comprising contacting a target nucleic acid with at least one
 PT reagent that distinguishes between methylated and non-methylated CpG
 PT dinucleotides.

PS Claim 26; SEQ ID NO 192; 448pp; English.

CC The invention relates to a method of detecting and differentiating
 CC between lymphoid cell proliferative disorders associated with at least
 CC one gene and/or their regulatory regions in a subject by contacting a
 CC target nucleic acid in a biological sample obtained from the subject with
 CC at least one reagent or series of reagents that distinguish between
 CC methylated and non-methylated CpG dinucleotides within the target nucleic
 CC acid. The genes and/or their regulatory regions are preferably selected
 CC from MBR1, CSNK2B, BGR4, AR, CDK4, RB2, CDC25A, GRB B2A, MYO1, CDH3,
 CC MYC1, ELK1, ABL1, APC, BC12, CDH1, CDKN1A, CDKN1B, CDKN2A, CDKN2B, FOS,
 CC GSTR1, HIC-1, MGMT, MLH1, MOS, MYC, PTEN, RBL2, TGFBR2, TP73, CDKN1C,
 CC GSK3beta, ESR1, APAF1, BAK1, BAX or HMOX5. Oligomers, peptide nucleic
 CC acid (PNA)-oligomers and/or isolated nucleic acids based on the sequences
 CC of the genes are useful for detecting the methylation state of all the
 CC CpG dinucleotides within one or more the sequences, or their complements,
 CC for determining the cytosine methylation state and or single nucleotide
 CC polymorphisms (SNPs), and for differentiating at least two of the medical
 CC conditions such as diffuse large B-cell lymphoma, mantle cell lymphoma,
 CC chronic lymphocytic leukemia, small lymphocytic lymphoma and follicular
 CC lymphoma. They are also useful for detecting of a predisposition to,
 CC differentiation between subtypes, diagnosis, prognosis, treating and/or
 CC monitoring of lymphoid cell proliferative disorder. This sequence
 CC represents a nucleic acid of a pretreated genomic DNA derived from the
 CC above mentioned genes.

XX Sequence 7110 BP; 1818 A; 0 C; 1721 G; 3571 T; 0 U; 0 Other;

Query Match 54.0%; Score 21.6; DB 10; Length 7110;
 Best Local Similarity 85.7%; Pred. No. 1.5e+02;

Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 ATACCTACTCCCAACATCATCTTAC 31
 DB 3083 AACCTAACCCCAACATCATCTTAC 3056

RESULT 14
 ADS89580/C
 ID ADS89580 standard; DNA; 7110 BP.

AC ADS89580;

XX 18-NOV-2004 (first entry)

DE Oligonucleotide of the invention SEQ ID NO:596.

XX ss; cell proliferative disorder; breast; methylation; cytostatic;
 KM gene therapy; single nucleotide polymorphism; SNP.

OS Unidentified.

XX WO2004035803-A2.

PD 29-APR-2004.

PF 01-OCT-2003; 2003WO-EP010881.

XX 01-OCT-2002; 2002DE-01045779.

PR 07-JAN-2003; 2003DE-01000096.

PR 17-APR-2003; 2003DE-01017955.

XX (EPIC-) EPICGENOMICS AG.

PI Foekens J, Harbeck N, Koenig T, Maier S, Martens J, Model F;
 PI Mmmrich I, Rujan T, Schmitt A, Schmitt M, Look MP, Marx A;

XX WPI, 2004-348468/32.

PT Predicting responsiveness of a subject with breast cell proliferative
 PT disorder, useful for treating or differentiating breast cell

PT proliferative disorders comprises analyzing methylation pattern of a
 PT genomic DNA from the subject.

PS Disclosure; SEQ ID NO 596; 104bp; English.

XX The invention relates to a novel method for predicting the responsiveness
 CC of a subject with a cell proliferative disorder of the breast tissues to
 CC a therapy comprising analysing the methylation pattern of a target
 CC nucleic acid by contacting at least one of the target nucleic acids in a
 CC biological sample obtained from the subject prior to or during treatment.
 CC The method of the invention has cytosolic activity, and may have a use
 CC in gene therapy. The set of oligonucleotides comprising at least two of
 CC the oligomers are useful for detecting the cytosine methylation state
 CC and/or single nucleotide polymorphisms (SNPs) within the sequences. The
 CC methods, nucleic acid, oligonucleotide, and kit are useful for the
 CC treatment, characterisation, classification and/or differentiation, of
 CC breast cell proliferative disorders. The method is also useful for
 CC predicting the responsiveness of a subject with a cell proliferative
 CC disorder of the breast tissues to a therapy. The present sequence is used
 CC in the exemplification of the invention.

XX Sequence 7110 BP; 1818 A; 0 C; 1721 G; 3571 T; 0 U; 0 Other;

Query Match 54.0%; Score 21.6; DB 13; Length 7110;
 Best Local Similarity 85.7%; Pred. No. 1.5e+02;

Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 ATACCTACTCCCAACATCATCTTAC 31
 DB 3083 AACCTAACCCCAACATCATCTTAC 3056

RESULT 15

ABZ13750/C
 ID ABZ13750 standard; DNA; 1341 BP.

XX AC ABZ13750;

DT 21-JAN-2003 (first entry)

DE Arabidopsis thaliana stress regulated gene SEQ ID NO 1555.

XX Arabidopsis thaliana; plant; gene; stress; transgenic; de.

XX Arabidopsis thaliana.

OS WO200216655-A2.

XX 28-FEB-2002.

PF 24-AUG-2001; 2001WO-US026685.

XX 24-AUG-2000; 2000US-0227866P.

PR 26-JAN-2001; 2001US-0264647P.

PR 22-JUN-2001; 2001US-0300111P.

XX (SCRI) SCRIPPS RES INST
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.

PI Harper JF, Kreps J, Wang X, Zhu T;

XX WPI; 2002-304127/34.

PT Identifying a stress condition to which a plant cell has been exposed and
 PT producing plants with increased tolerance to these abiotic stresses.

XX Claim 14; SEQ ID NO 1555; 577bp + Sequence Listing; English.

XX The invention relates to identifying a stress condition to which a plant
 CC cell has been exposed, comprising: (a) contacting nucleic acid
 CC representative of expressed polynucleotides in the plant cell with an
 CC array or probes representative of the plant cell genome; and (b)
 CC detecting a profile of expressed polynucleotides in the plant cell
 CC characteristic of a stress response. The method is useful in the
 CC production of transgenic plants, cells and seeds and in producing plants
 CC with increased tolerance to abiotic stress. The present sequence is that
 CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
 CC in methods of the invention. Note: The sequence data for this patent is
 CC not represented in the printed specification but is based on sequence
 CC information supplied to Derwent by the European Patent Office

XX Sequence 1341 BP; 460 A; 231 C; 319 G; 331 T; 0 U; 0 Other;

Query Match 53.5%; Score 21.4; DB 6; Length 1341;
 Best Local Similarity 71.8%; Pred. No. 1.5e+02;

Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 AGATATCTACTCCCAACATCATCTTACTGACAA 39
 DB 288 AGTCCACCTATTCACCAACATCTCTGTCTCACACAA 250

Search completed: September 15, 2005, 21:33:23
 Job time : 20.3656 secs

This Page Blank (uspto)

Sat Sep 17 09:06:49 2005

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using bw model

Run on: September 14, 2005, 19:14:19 ; Search time 6.05048 Seconds
(without alignments)
10817.505 Million cell updates/sec

Title: US-10-713-381-2_COPY_1239_1278

Perfect score: 40
Sequence: 1 aggataccaccctcccaacaacacccatccatccatgcgaac 40Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/prodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq: *
6: /cgn2_6/prodata/1/ina/backfile601.seq: *Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	1394	3	US-08-880-499-1
2	40	100.0	1394	4	US-08-880-499-2
3	20.6	51.5	1071	4	US-09-248-796A-5546
4	20.6	51.5	42075	4	US-09-949-016-14895
5	20.4	51.0	601	4	US-09-949-016-65770
6	20.4	51.0	973	4	US-09-774-528-184
7	20.4	51.0	4201	3	US-08-945-056-4
8	20.4	51.0	51770	4	US-09-949-016-13668
9	20.4	51.0	84571	4	US-09-949-016-17420
10	20.4	51.0	126200	4	US-09-949-016-11824
11	20.4	51.0	126200	4	US-09-949-016-13193
12	20.2	50.5	5099	3	US-09-610-040-5
13	20.2	50.5	5099	4	US-10-267-763-5
14	20	50.0	601	4	US-09-949-016-88655
15	20	50.0	2615	4	US-09-620-312D-22
16	20	50.0	77586	4	US-09-949-016-13220
17	20	50.0	77586	4	US-09-949-016-13221
18	19.8	49.5	107	4	US-09-513-999C-19002
19	19.8	49.5	601	4	US-09-949-016-202698
20	19.8	49.5	1017	4	US-09-710-279-2867
21	19.8	49.5	1164	3	US-09-134-001C-2534
22	19.8	49.5	2066	4	US-09-248-796A-6306
23	19.8	49.5	2971	4	US-09-710-279-3449
24	19.8	49.5	9366	4	US-09-949-016-17437
25	19.8	49.5	74097	4	US-09-949-016-11785
26	19.8	49.5	74097	4	US-09-949-016-16239
27	19.8	49.5	265038	4	US-09-949-016-15779

28	19.6	49.0	189	4	US-09-540-236-1562	Sequence 1562, Ap
29	19.6	49.0	601	4	US-09-949-016-83110	Sequence 83110, A
30	19.6	49.0	601	4	US-09-949-016-83111	Sequence 83111, A
31	19.6	49.0	4322	2	US-08-537-342-1	Sequence 1, Appl
32	19.6	49.0	6330	4	US-09-949-016-15506	Sequence 15506, A
33	19.6	49.0	66986	4	US-09-596-002-29	Sequence 29, Appl
34	19.6	49.0	75480	4	US-09-949-016-16090	Sequence 16090, A
35	19.6	49.0	80858	4	US-09-949-016-12859	Sequence 12859, A
36	19.6	49.0	80859	4	US-09-949-016-15715	Sequence 15715, A
37	19.6	49.0	89892	4	US-09-949-016-13667	Sequence 13667, A
38	19.6	49.0	107937	4	US-09-949-016-17192	Sequence 17192, A
39	19.6	49.0	140315	4	US-09-949-016-14141	Sequence 14141, A
40	19.4	48.5	601	4	US-09-949-016-26227	Sequence 26227, A
41	19.4	48.5	601	4	US-09-949-016-74129	Sequence 74129, A
42	19.4	48.5	601	4	US-09-949-016-74160	Sequence 74160, A
43	19.4	48.5	601	4	US-09-949-016-202761	Sequence 202761, A
44	19.4	48.5	2810	4	US-09-949-016-5447	Sequence 5447, Ap
45	19.4	48.5	2816	3	US-09-643-597-333	Sequence 333, Ap

ALIGNMENTS

RESULT 1
US-08-880-499-1
Sequence 1, Application US/08880499

Patent No. 6037523

GENERAL INFORMATION:

APPLICANT: Alberson, Marc C.

APPLICANT: Fox, Tim W.

APPLICANT: Carl, Garnat W.

APPLICANT: Hultman, Gary A.

APPLICANT: Kendall, Timmy L.

TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSES: PIONEER HI-BRED INTERNATIONAL, INC.

STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.

STREET: Box 1000

CITY: Johnston

STATE: Iowa

COUNTRY: USA

ZIP: 50131

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/880,499

FILING DATE: CONCURRENTLY HEREWITH

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Sweeney, Patricia A.

REGISTRATION NUMBER: 32,733

REFERENCE/DOCKET NUMBER: 0578

TELECOMMUNICATION INFORMATION:

TELEPHONE: (515) 248-4800

TELEFAX: (515) 248-4844

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1394 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-880-499-1

Query Match 100.0%; Score 40; DB 3; Length 1394;
Best Local Similarity 100.0%; Pred. No. 4.6e-07;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCGAACATCATCTTACTCATGCAAC 40
|||||
Db 1239 AGGATACCTACTCCGAACATCATCTTACTCATGCAAC 1278

RESULT 2

US-08-880-499-2
Sequence 2, Application US/0880499
Patent No. 6037523
GENERAL INFORMATION:
APPLICANT: Albertson, Marc C.
APPLICANT: Fox, Tim W.
APPLICANT: Carl, Garnat W.
APPLICANT: Huffman, Gary A.
APPLICANT: Kendall, Timmy L.
TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION
TITLE OF INVENTION: AND METHOD OF USING SAME
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.
STREET: Box 1000
CITY: Johnston
STATE: Iowa
COUNTRY: USA
ZIP: 50131
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/880,499
FILING DATE: CONCURRENTLY HERewith
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Sweeney, Patricia A.
REGISTRATION NUMBER: 32,733
REFERENCE/DOCKET NUMBER: 0578
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 248-4800
TELEFAX: (515) 248-4844
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1394 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-880-499-2

Query Match 100.0%; Score 40; DB 3; Length 1394;

Best Local Similarity 100.0%; Pred. No. 4,6e-07;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCGAACATCATCTTACTCATGCAAC 40
|||||
Db 1239 AGGATACCTACTCCGAACATCATCTTACTCATGCAAC 1278

RESULT 3

US-09-248-796A-5546
Sequence 5546, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstock et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196,132
CURRENT APPLICATION NUMBER: US/09/248,796A
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13

PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 5546
LENGTH: 1071
TYPE: DNA
ORGANISM: Candida albicans
US-09-248-796A-5546

Query Match 51.5%; Score 20.6; DB 4; Length 1071;

Best Local Similarity 74.3%; Pred. No. 51;
Matches 26; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 4 ATACCTACTCCGAACATCATCTTACTCATGCA 38
|||||
Db 224 ATCCCACTCTCAACATGTCACATCTTACTTCA 258

RESULT 4

US-09-949-016-14995/C
Sequence 14995, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14995
LENGTH: 42075
TYPE: DNA
ORGANISM: Human
US-09-949-016-14995

Query Match 51.5%; Score 20.6; DB 4; Length 42075;

Best Local Similarity 85.2%; Pred. No. 1.2e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 ATACCTACTCCGAACATCATCTTCA 30
|||||
Db 31591 ATCCCTACCACAAACATCATCTTCA 31565

RESULT 5

US-09-949-016-65770/C
Sequence 65770, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 65770
LENGTH: 601

TYPE: DNA
ORGANISM: Human
US-09-949-016-65770

Query Match
Best Local Similarity 51.0%; Score 20.4; DB 4; Length 601;
Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 GGATCTACTCCCAACATCCATCTTACTCATGCAA 39
DB 414 GGATCTCTGGCTCAGCATCCTCTGCTCAGGCAA 377

RESULT 6
US-09-774-528-184
Sequence 184, Application US/09774528

PATENT No. 6743619
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Zhou, Ping
APPLICANT: Goodrich, Kyle
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Ren, Feiyun
APPLICANT: Zhang, Jie
APPLICANT: Zhao, Qing A.
APPLICANT: Yang, Yonghong
APPLICANT: Xue, Aidong J.
APPLICANT: Wehrman, Tom
APPLICANT: Wang, Jian-Rui
APPLICANT: Wang, Dunrui
APPLICANT: Dmanac, Radoje T.
TITLE OF INVENTION: No. 6743619e1 Nucleic Acids and
FILE REFERENCE: 802
CURRENT APPLICATION NUMBER: US/09/774,528
CURRENT FILING DATE: 2001-01-30
NUMBER OF SEQ ID NOS: 441
SOFTWARE: pf_fl_genes Version 2.0
SEQ ID NO 184
LENGTH: 973
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (231)..(749)
US-09-774-528-184

Query Match
Best Local Similarity 51.0%; Score 20.4; DB 4; Length 973;
Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 GGATCTACTCCCAACATCCATCTTACTCATGCAA 39
DB 626 GGAATCACTCCAAAAGAAACCTCGAACAATGCCA 663

RESULT 7
US-08-945-056-4
Sequence 4, Application US/08945056
PATENT No. 6077994
GENERAL INFORMATION:
APPLICANT: Coupland, George M.
APPLICANT: Putterill, Joanna J.
TITLE OF INVENTION: Genetic control of flowering
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSER: Nixon & Vanderhye PC
STREET: 8th Floor, 1100 No. 6077994th Glebe Road
CITY: Arlington
STATE: Virginia
COUNTRY: United States of America
COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/945,056
FILING DATE: 20-OCT-1997
CLASSIFICATION: 800

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB95/02561
FILING DATE: 01-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9422083.7
FILING DATE: 02-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ms Mary J Wilson
REGISTRATION NUMBER: 32,955
REFERENCE/DOCKET NUMBER: 620-17
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 4201 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Arabidopsis thaliana
STRAIN: Landsberg erecta
POSITION IN GENOME:
MAP POSITION: chromosome 5
US-08-945-056-4

Query Match
Best Local Similarity 51.0%; Score 20.4; DB 3; Length 4201;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 GATACCTACTCCCAACATCCATCTTACT 32
DB 2631 GATACGAGCTCCACACCATCAACTTACT 2660

RESULT 8
US-09-949-016-13668/c
Sequence 13668, Application US/09949016
PATENT No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13668
LENGTH: 51770
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(51770)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13668

Query Match
51.0%; Score 20.4; DB 4; Length 51770;

Best Local Similarity 71.1%; Pred. No. 1.5e+02;
Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 2 GGATACCTACTCCCAACAATCCATCTTACTCATGCAA 39

Db 9851 GGATTCCTGGGCTCAAGCAATCCTCCTGCTCAGGCAA 9814

RESULT 9
US-09-949-016-17420

/ Sequence 17420, Application US/09949016

/ Patent No. 6812339

/ GENERAL INFORMATION:

/ APPLICANT: VENTER, J. Craig et al.

/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

/ WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

/ FILE REFERENCE: CL001307

/ CURRENT APPLICATION NUMBER: US/09/949,016

/ CURRENT FILING DATE: 2000-04-14

/ PRIOR APPLICATION NUMBER: 60/241,755

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR APPLICATION NUMBER: 60/237,768

/ PRIOR FILING DATE: 2000-10-03

/ PRIOR APPLICATION NUMBER: 60/231,498

/ PRIOR FILING DATE: 2000-09-08

/ NUMBER OF SEQ ID NOS: 207012

/ SOFTWARE: FastSeq for Windows Version 4.0

/ SEQ ID NO 17420

/ LENGTH: 84571

/ TYPE: DNA

/ ORGANISM: Human

/ NAME/KEY: misc_feature

/ LOCATION: (1)...(84571)

/ OTHER INFORMATION: n = A,T,C or G

/ US-09-949-016-17420

Query Match 51.0%; Score 20.4; DB 4; Length 84571;
Best Local Similarity 71.1%; Pred. No. 1.7e+02;

Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 2 GGATACCTACTCCCAACAATCCATCTTACTCATGCAA 39

Db 24559 GGAAATCAACTCCAAAGAAACCTTCAAAACCATGCAA 24596

RESULT 10
US-09-949-016-11824/C

/ Sequence 11824, Application US/09949016

/ Patent No. 6812339

/ GENERAL INFORMATION:

/ APPLICANT: VENTER, J. Craig et al.

/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

/ WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

/ FILE REFERENCE: CL001307

/ CURRENT APPLICATION NUMBER: US/09/949,016

/ CURRENT FILING DATE: 2000-04-14

/ PRIOR APPLICATION NUMBER: 60/241,755

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR APPLICATION NUMBER: 60/237,768

/ PRIOR FILING DATE: 2000-10-03

/ PRIOR APPLICATION NUMBER: 60/231,498

/ PRIOR FILING DATE: 2000-09-08

/ NUMBER OF SEQ ID NOS: 207012

/ SOFTWARE: FastSeq for Windows Version 4.0

/ SEQ ID NO 11824

/ LENGTH: 126200

/ TYPE: DNA

/ ORGANISM: Human

/ NAME/KEY: misc_feature

/ LOCATION: (1)...(126200)

/ OTHER INFORMATION: n = A,T,C or G

US-09-949-016-11824

Query Match 51.0%; Score 20.4; DB 4; Length 126200;
Best Local Similarity 71.1%; Pred. No. 1.9e+02;

Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 2 GGATACCTACTCCCAACAATCCATCTTACTCATGCAA 39

Db 53133 GGAAATCAACTCCAAAGAAACCTTCAAAATCATGCAA 53096

RESULT 11
US-09-949-016-13193/C

/ Sequence 13193, Application US/09949016

/ Patent No. 6812339

/ GENERAL INFORMATION:

/ APPLICANT: VENTER, J. Craig et al.

/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

/ WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

/ FILE REFERENCE: CL001307

/ CURRENT APPLICATION NUMBER: US/09/949,016

/ CURRENT FILING DATE: 2000-04-14

/ PRIOR APPLICATION NUMBER: 60/241,755

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR APPLICATION NUMBER: 60/237,768

/ PRIOR FILING DATE: 2000-10-03

/ PRIOR APPLICATION NUMBER: 60/231,498

/ PRIOR FILING DATE: 2000-09-08

/ NUMBER OF SEQ ID NOS: 207012

/ SOFTWARE: FastSeq for Windows Version 4.0

/ SEQ ID NO 13193

/ LENGTH: 126200

/ TYPE: DNA

/ ORGANISM: Human

/ NAME/KEY: misc_feature

/ LOCATION: (1)...(126200)

/ OTHER INFORMATION: n = A,T,C or G

/ US-09-949-016-13193

Query Match 51.0%; Score 20.4; DB 4; Length 126200;
Best Local Similarity 71.1%; Pred. No. 1.9e+02;

Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 2 GGATACCTACTCCCAACAATCCATCTTACTCATGCAA 39

Db 53133 GGAAATCAACTCCAAAGAAACCTTCAAAATCATGCAA 53096

RESULT 12
US-09-610-040-5

/ Sequence 5, Application US/09610040

/ Patent No. 6465217

/ GENERAL INFORMATION:

/ APPLICANT: Boyes, Douglas

/ APPLICANT: Davis, Keith

/ APPLICANT: Woessner, Jeffrey

/ APPLICANT: Gorlach, Jorn

/ APPLICANT: Hamilton, Carol

/ APPLICANT: Hoffman, Neil

/ APPLICANT: Kloci, Andreas

/ APPLICANT: Zayed, Adel

/ APPLICANT: Ascenzi, Robert

/ TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF CHORISMATE SYNTHAS

/ WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

/ FILE REFERENCE: 9128.14

/ CURRENT APPLICATION NUMBER: US/09/610,040

/ CURRENT FILING DATE: 2000-07-05

/ NUMBER OF SEQ ID NOS: 11

/ SOFTWARE: PatentIn version 3.0

/ SEQ ID NO 5

/ LENGTH: 5099

/ TYPE: DNA

```
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (4049)..(4049)
; OTHER INFORMATION: "n" indicates any nucleotide (A, C, T, or G).
US-09-610-040-5

Query Match          50.5%; Score 20.2; DB 3; Length 5099;
Best Local Similarity 88.0%; Pred. No. 1.1e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      6  ACCTACTCCCAACATCCATCTTA 30
DB      2855  ACCTACCCCAACATCCCTTTTA 2879

RESULT 13
US-10-267-763-5
; Sequence 5, Application US/10267763
; Patent No. 6800459
; GENERAL INFORMATION:
; APPLICANT: Goriach, Jörn
; APPLICANT: Boyes, Douglas
; APPLICANT: Davis, Keith
; APPLICANT: Hamilton, Carol
; APPLICANT: Hoffman, Neil
; APPLICANT: Klotz, Andreas
; APPLICANT: Woessner, Jeffrey
; APPLICANT: Zayed, Adel
; APPLICANT: Ascenzi, Robert
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
; TITLE OF INVENTION: MODULATION OF CHORISMATE SYNTHASE AND CHORISMATE MUTASE
; TITLE OF INVENTION: EXPRESSION OR ACTIVITY IN PLANTS
; FILE REFERENCE: 2035DIV1
; CURRENT APPLICATION NUMBER: US/10/267,763
; CURRENT FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 09/610,040
; PRIOR FILING DATE: 2000-07-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 5099
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(5099)
; OTHER INFORMATION: n = A,T,C or G
US-10-267-763-5

Query Match          50.5%; Score 20.2; DB 4; Length 5099;
Best Local Similarity 88.0%; Pred. No. 1.1e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      6  ACCTACTCCCAACATCCATCTTA 30
DB      2855  ACCTACCCCAACATCCCTTTTA 2879

RESULT 14
US-09-949-016-88655
; Sequence 88655, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
```

```
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 88655
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-88655

Query Match          50.0%; Score 20; DB 4; Length 601;
Best Local Similarity 68.4%; Pred. No. 79;
Matches 26; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY      2  GGATACCTACTCCCAACATCCATCTTACTCATGCA 39
DB      293  GTATAGTCTCCCAACGCTGCTTCTTATATATAAA 330

RESULT 15
US-09-620-312D-22
; Sequence 22, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aldong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunging
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhilwei
; APPLICANT: John Tillinghaest
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 22
; LENGTH: 2615
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (160)..(849)
US-09-620-312D-22

Query Match          50.0%; Score 20; DB 4; Length 2615;
Best Local Similarity 72.2%; Pred. No. 1.1e+02;
Matches 26; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY      4  AATACCTACTCCCAACATCCATCTTACTCATGCA 39
DB      1855  ATACCAATTCCCAACCAACATCTTCCATTGGAA 1890

Search completed: September 15, 2005, 08:25:23
Job time : 7.05048 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 07:10:11 ; Search time 26.0505 Seconds
(without alignments)
10230.248 Million cell updates/sec

Title: US-10-713-381-2_COPY_1239_1278

Perfect score: 40
Sequence: 1 aggatacctaccatcccaaacatccatctaccatgcgaac 40

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7389322 seqs, 333128559 residues

Total number of hits satisfying chosen parameters: 14778644

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
22: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
23: /cgn2_6/ptodata/2/pubpna/US11_PUBCOMB.seq:*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	100.0	40	US-10-713-381-6	Sequence 6, Appli
2	40	100.0	50	US-10-713-381-5	Sequence 5, Appli
3	40	100.0	158	US-10-713-381-3	Sequence 3, Appli
4	40	100.0	255	US-10-713-381-9	Sequence 9, Appli
5	40	100.0	1394	US-10-713-381-1	Sequence 1, Appli
6	40	100.0	1394	US-10-713-381-2	Sequence 2, Appli
7	24	60.0	85779	US-10-713-993-6880	Sequence 6880, Ap

8	23.2	58.0	706	18	US-10-424-599-31006	Sequence 31006, A
9	22.6	56.5	294	20	US-10-425-115-143168	Sequence 143168, A
10	22.4	56.0	315	18	US-10-424-599-83179	Sequence 83179, A
11	22.4	56.0	226215	13	US-10-087-192-1948	Sequence 1948, A
12	21.6	54.0	4110	20	US-10-473-126-342	Sequence 342, App
13	21.6	54.0	6309	15	US-10-311-455-277	Sequence 277, App
14	21.4	53.5	1341	9	US-09-938-842A-1555	Sequence 1555, App
15	21.4	53.5	1341	11	US-09-938-842A-1555	Sequence 1555, App
16	21.4	53.5	6881	15	US-10-311-455-128	Sequence 128, App
17	21.4	53.5	6881	16	US-10-240-452-4	Sequence 4, Appli
18	21.4	53.5	8168	15	US-10-311-455-901	Sequence 901, App
19	21.4	53.5	8168	16	US-10-240-454-23	Sequence 23, Appli
20	21.4	53.5	3673778	16	US-10-312-841-1	Sequence 1, Appli
21	21.4	53.5	3673778	16	US-10-312-841-1	Sequence 1, Appli
22	21	52.5	293	18	US-10-424-599-24043	Sequence 24043, A
23	21	52.5	931	18	US-10-424-599-20320	Sequence 20320, A
24	21	52.5	2000	17	US-10-260-238-2594	Sequence 2594, App
25	21	52.5	2792	18	US-10-041-018-361	Sequence 361, App
26	21	52.5	3117	18	US-10-041-018-362	Sequence 362, App
27	21	52.5	5507	20	US-10-473-126-198	Sequence 198, App
28	21	52.5	5507	20	US-10-473-126-198	Sequence 344, App
29	21	52.5	8404	15	US-10-311-455-1568	Sequence 1568, App
30	21	52.5	8404	18	US-10-221-714A-222	Sequence 222, App
31	21	52.5	10286	14	US-10-239-676-13	Sequence 13, Appli
32	21	52.5	10286	15	US-10-240-453-21	Sequence 21, Appli
33	21	52.5	63294	11	US-09-997-722-205	Sequence 205, App
34	21	52.5	106315	13	US-10-087-192-178	Sequence 178, App
35	21	52.5	2140405	13	US-10-027-632-76212	Sequence 76212, A
36	21	52.5	2140405	17	US-10-027-632-76212	Sequence 76212, A
37	20.8	52.0	3263	19	US-10-437-963-4616	Sequence 4616, App
38	20.8	52.0	3292	13	US-10-027-632-112652	Sequence 112652, App
39	20.8	52.0	3292	17	US-10-027-632-112652	Sequence 112652, App
40	20.8	52.0	6620	14	US-10-239-676-196	Sequence 196, App
41	20.8	52.0	6620	15	US-10-240-453-290	Sequence 290, App
42	20.8	52.0	513309	10	US-09-754-853A-4	Sequence 4, Appli
43	20.6	51.5	970	18	US-10-424-599-91073	Sequence 91073, A
44	20.6	51.5	1103	22	US-10-499-352A-170	Sequence 170, App
45	20.6	51.5	1889	22	US-10-499-352A-178	Sequence 178, App

ALIGNMENTS

RESULT 1
US-10-713-381-6
; Sequence 6, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFMAN, GARY L.
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 576R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 6
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-6
Query Match 100.0%; Score 40; DB 20; Length 40;
Best Local Similarity 100.0%; Pred. No. 3.2e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGATACCTACCTCCCAACATCCATCTTACATGACAC 40

Db 1 AGGATACCTACTCCCAAGATCATCTTACTCATGCAAC 40

RESULT 2
US-10-713-381-5
; Sequence 5, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-5

Query Match 100.0%; Score 40; DB 20; Length 50;
Best Local Similarity 100.0%; Pred. No. 3.3e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAAGATCATCTTACTCATGCAAC 40
Db 11 AGGATACCTACTCCCAAGATCATCTTACTCATGCAAC 50

RESULT 3
US-10-713-381-3
; Sequence 3, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 158
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-3

Query Match 100.0%; Score 40; DB 20; Length 158;
Best Local Similarity 100.0%; Pred. No. 4.2e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAAGATCATCTTACTCATGCAAC 40
Db 86 AGGATACCTACTCCCAAGATCATCTTACTCATGCAAC 125

RESULT 4

US-10-713-381-9
; Sequence 9, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 255
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-9

Query Match 100.0%; Score 40; DB 20; Length 255;
Best Local Similarity 100.0%; Pred. No. 4.6e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAAGATCATCTTACTCATGCAAC 40
Db 99 AGGATACCTACTCCCAAGATCATCTTACTCATGCAAC 138

RESULT 5
US-10-713-381-1
; Sequence 1, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-1

Query Match 100.0%; Score 40; DB 20; Length 1394;
Best Local Similarity 100.0%; Pred. No. 6.4e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAAGATCATCTTACTCATGCAAC 40
Db 1239 AGGATACCTACTCCCAAGATCATCTTACTCATGCAAC 1278

RESULT 6
US-10-713-381-2
; Sequence 2, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.

```
/ APPLICANT: FOX, TIMOTHY W.
/ APPLICANT: GARNAT, CARL W.
/ APPLICANT: HUFMAN, GARY
/ APPLICANT: KENDALL, TIMMY L.
/ TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
/ FILE REFERENCE: 578R
/ CURRENT APPLICATION NUMBER: US/10/713,381
/ CURRENT FILING DATE: 2003-11-14
/ PRIOR APPLICATION NUMBER: 08/880,499
/ PRIOR FILING DATE: 1997-06-23
/ NUMBER OF SEQ ID NOS: 24
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 2
/ LENGTH: 1394
/ TYPE: DNA
/ ORGANISM: Zea mays
US-10-713-381-2

Query Match          100.0%; Score 40; DB 20; Length 1394;
Best Local Similarity 100.0%; Pred. No. 6,4e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AGGATACCTACTCCCAACATCATCTTACTCATGCAAC 40
DB 1239 AGGATACCTACTCCCAACATCATCTTACTCATGCAAC 1278

RESULT 7
US-10-719-993-6880/C
/ Sequence 6880, Application US/10719993
/ Publication No. US20040265849A1
/ GENERAL INFORMATION:
/ APPLICANT: CARGILL, Michele et al.
/ TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
/ FILE REFERENCE: CL001496
/ CURRENT APPLICATION NUMBER: US/10/719,993
/ CURRENT FILING DATE: 2003-11-24
/ NUMBER OF SEQ ID NOS: 55342
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 6880
/ LENGTH: 85779
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1)..(85779)
/ OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-719-993-6880

Query Match          60.0%; Score 24; DB 20; Length 85779;
Best Local Similarity 84.4%; Pred. No. 38;
Matches 27; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 4 ATACCTACTCCCAACATCATCTTACTCAT 35
DB 16082 AAACCTACTCCCAACATCATCTTACTCAT 16051

RESULT 8
US-10-424-599-31006
/ Sequence 31006, Application US/10424599
/ Publication No. US20040031072A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J
/ APPLICANT: Kovalic, David K
/ APPLICANT: Zhou Yihua
/ APPLICANT: Cao Yongwei
/ TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(53223)B
/ CURRENT APPLICATION NUMBER: US/10/424,599
/ NUMBER OF SEQ ID NOS: 285684
/ SEQ ID NO 31006
/ LENGTH: 315
/ TYPE: DNA
/ ORGANISM: Glycine max
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT3847_46127C.1
US-10-424-599-83179
```

```
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 285684
/ SEQ ID NO 31006
/ LENGTH: 706
/ TYPE: DNA
/ ORGANISM: Glycine max
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT3847_128002C.1
US-10-424-599-31006

Query Match          58.0%; Score 23.2; DB 18; Length 706;
Best Local Similarity 77.8%; Pred. No. 31;
Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 5 TACCTACTCCCAACATCATCTTACTCATGCAAC 40
DB 455 TCCCTTCTCCCAACCCCTCATCAACCTCATCAAC 490

RESULT 9
US-10-425-115-143168
/ Sequence 143168, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(53222)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 143168
/ LENGTH: 294
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MRT4577_62049C.1
US-10-425-115-143168

Query Match          56.5%; Score 22.6; DB 20; Length 294;
Best Local Similarity 86.2%; Pred. No. 45;
Matches 25; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 4 ATACCTACTCCCAACATCATCTTACT 32
DB 226 ATACCTGTCCCATACCTCATCTTACT 254

RESULT 10
US-10-424-599-83179
/ Sequence 83179, Application US/10424599
/ Publication No. US20040031072A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J
/ APPLICANT: Kovalic, David K
/ APPLICANT: Zhou Yihua
/ APPLICANT: Cao Yongwei
/ TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(53223)B
/ CURRENT APPLICATION NUMBER: US/10/424,599
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 285684
/ SEQ ID NO 83179
/ LENGTH: 315
/ TYPE: DNA
/ ORGANISM: Glycine max
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT3847_46127C.1
US-10-424-599-83179
```

Query Match 56.0%; Score 22.4; DB 18; Length 315;
Best Local Similarity 81.2%; Pred. No. 55;
Matches 26; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 ATACCTACTCCCAACATCATCTTACTCAT 35
DB 211 AACGAATCTCTTAACATCATCTTAAATCCT 242

RESULT 11
US-10-087-192-1948
; Sequence 1948, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1948
; LENGTH: 226215
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-1948

Query Match 56.0%; Score 22.4; DB 13; Length 226215;
Best Local Similarity 72.5%; Pred. No. 2e+02;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAACATCATCTTACTCATGCAAC 40
DB 183206 AGAATTCCTGCTTCCAAACACAGCATCTTACTCTTTTAC 183245

RESULT 12
US-10-473-126-342/C
; Sequence 342, Application US/10473126
; Publication No. US20040234973A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Methods and nucleic acids for the analysis of hematopoietic cell
; FILE REFERENCE: proliferative disorders
; CURRENT APPLICATION NUMBER: US/10/473,126
; PRIOR FILING DATE: 2003-09-26
; NUMBER OF SEQ ID NOS: 1258
; SEQ ID NO 342
; LENGTH: 4110
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-473-126-342

Query Match 54.0%; Score 21.6; DB 20; Length 4110;
Best Local Similarity 85.7%; Pred. No. 1.9e+02;
Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 ATACCTACTCCCAACATCATCTTACTCAT 31
DB 3083 AAACCTAACCCCAACATCATCTTAC 3056

RESULT 13

US-10-311-455-277/C
; Sequence 277, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 277
; LENGTH: 6309
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-277

Query Match 54.0%; Score 21.6; DB 15; Length 6309;
Best Local Similarity 75.0%; Pred. No. 2.1e+02;
Matches 27; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 4 ATACCTACTCCCAACATCATCTTACTCATGCAA 39
DB 3335 ATACCTATTCCTATCCAGACATCTCATATATCCAA 3300

RESULT 14
US-09-938-842A-1555/C
; Sequence 1555, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1555
; LENGTH: 1341
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1555

Query Match 53.5%; Score 21.4; DB 9; Length 1341;
Best Local Similarity 71.8%; Pred. No. 1.9e+02;
Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAACATCATCTTACTCATGCAA 39
DB 288 AGTCACCTATTCCTCCAAACATCTCTGTGCTTACACAA 250

RESULT 15


```
US-09-938-842A-1555/c
; Sequence 1555, Application US/09938842A
; Publication No. US2004009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938, 842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1555
; LENGTH: 1341
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1555

Query Match      53.5%; Score 21.4; DB 11; Length 1341;
Best Local Similarity 71.8%; Pred. No. 1.9e+02;
Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY      1 AGGATACCTACTCCCAACATCCATCTTACTCATGCAA 39
DB      288 AGTCCACCTATTCCTCAACATCTCTGTGCTCAGACAA 250

Search completed: September 15, 2005, 20:46:11
Job time : 32.0505 sec
```

This page blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 16:07:32 ; Search time 124.754 Seconds
(without alignments)
12204.542 Million cell updates/sec

Title: US-10-713-381-2_COPY_1239_1278

Perfect score: 40
Sequence: 1 aggataccctaccctcccaacacatccatccatccatgcgaac 40

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*

Pred. No. is the number of results predicted by a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	40	100.0	687	9	CC656939 OGMDO20TV
2	40	100.0	915	9	CG224225 OG1AG08TV
3	40	100.0	963	9	CC656933 OGMDO20TM
4	23.8	59.5	528	7	H24137 YMS0908.r1
5	23.6	59.0	375	8	B2706605 SM416B1-G
6	23.2	58.0	738	9	CO117589 CR181951
7	23.2	58.0	738	9	CR181951 Reverse s
8	23.2	58.0	1018	9	CL109241 ISB1-51D1
9	23.2	58.0	1253	6	CD495749 CDA18-D05
C 10	23	57.5	405	8	AZ883408 RPCI-23-1
C 11	23	57.5	798	4	BG506830 601861443
C 12	23	57.5	968	4	BF143726 601788606
13	22.8	57.0	771	4	BM406466 EST580781
14	22.6	56.5	410	2	AW257178 EST305315
15	22.6	56.5	560	2	BE205271 EST397947
C 16	22.6	56.5	624	9	CR843833 GR0AA78D
C 17	22.6	56.5	630	8	BZ173201 CH230-369
C 18	22.6	56.5	644	8	CR843905 GR0AA79A
19	22.6	56.5	664	8	AZ400686 IM0167P01
20	22.6	56.5	691	2	AM687128 NF006C08R
C 21	22.6	56.5	786	9	CR818130 GR0AA43D
C 22	22.4	56.0	1080	9	CL058859 CH216-88P
23	22.4	56.0	298	2	BB264116 BB264116
C 24	22.4	56.0	940	7	CK864514 AGENCOURT

C 25	22.4	56.0	940	7	CK871857 AGENCOURT
C 26	22.4	56.0	951	7	CN317130 AGENCOURT
27	22.2	55.5	273	1	AI206216 q27f06.x
28	22.2	55.5	346	1	AI698201 wa67911.x
29	22.2	55.5	398	1	AA725064 a106h07.s
30	22.2	55.5	417	1	AI467879 t178e09.x
31	22.2	55.5	479	2	BF601210 266128 MA
32	22.2	55.5	491	1	AI382034 c633905.x
C 33	22.2	55.5	609	4	BM027586 G1T000109
C 34	22.2	55.5	727	7	CK949745 4074916 B
C 35	22.2	55.5	802	9	CG811817 FSAA435TR
C 36	22.2	55.5	940	9	CG383910 OGZAP61TV
C 37	22.2	55.5	1090	9	CL084441 ISB1-3G16
C 38	22	55.0	207	7	CK913825 p3fmgcf_0
C 39	22	55.0	330	3	CNS098G7 BX071651 Single re
C 40	22	55.0	467	9	CG982704 CH240_164
C 41	22	55.0	478	3	CNS098GU BX047050 Single re
C 42	22	55.0	506	3	CNS09070 BX036328 Single re
C 43	22	55.0	506	8	AZ021739 RPCI-23-3
C 44	22	55.0	548	3	CNS09071 BX036328 Single re
C 45	22	55.0	576	8	AQ321478 RPCI11-10

ALIGNMENTS

RESULT 1	CC656939	687 bp	DNA	linear	GSS 19-JUN-2003
LOCUS	OGMDO20TV_ZM_0.7_1.5_KB_Zea_mays				
DEFINITION	genomic survey sequence.				
ACCESSION	CC656939				
VERSION	CC656939.1	GI:32060231			
KEYWORDS	GSS.				
SOURCE	Zea mays				
ORGANISM	Zea mays				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.				
AUTHORS	Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick,A., Fraser,C.M., Buddiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.				
TITLE	Consortium for Maize Genomics				
JOURNAL	Unpublished (2002)				
COMMENT	Other GSSs: OGMDO20TM Contact: Cathy Whitelaw TIGR				
FEATURES	9712 Medical Center Drive, Rockville, MD 20850, USA Tel: 301-838-5843 Fax: 301-838-0208 Email: whitelaw@tigr.org Seq primer: TF Class: sheared ends.				
source	Location/Qualifiers 1..687 /organism="Zea mays" /mol_type="genomic DNA" /strain="B73" /db_xref="taxon:4577" /clone="ZMMMA0554D15" /clone_11b="ZM_0.7_1.5_KB" /note="Vector: pBCSK-; Site 1: HincII, 0.7-1.5 kb methylation filtered genomic DNA library"				
ORIGIN	Query Match 100.0%; Score 40; DB 9; Length 687; Best Local Similarity 100.0%; Pred. No. 3.2e-05; Matches 40; Conservative 0; Mismatches 0; Indels 0; 1 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 40 337 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 298 				

RESULT 2
CG224225 915 bp DNA linear GSS 22-AUG-2003
LOCUS OGIAG08TV ZM 0.7.1.5 KB Zea mays genomic clone ZMMBMA0716B15,
DEFINITION genomic survey sequence.
ACCESSION CG224225
VERSION CG224225.1 GI:34124113
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 915)
Whiteaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other_GSSs: OGIAG08TH
Contact: Cathy Whiteaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteaw@tigr.org
Seq primer: TF
Class: sheared ends.
Location/Qualifiers
1..915
/organism="Zea mays"
/mol_type="genomic DNA"
/db_xref="taxon:4577"
/db_xref="B73"
/clone="ZMMBMA0716B15"
/clone_lib="ZM 0.7.1.5 KB"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match 100.0%; Score 40; DB 9; Length 915;
Best Local Similarity 100.0%; Pred. No. 3.4e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCTCCAAACATCCATCTTACTCATGCAAC 40
|||||
618 AGGATACCTACTCTCCAAACATCCATCTTACTCATGCAAC 657
|||||

RESULT 3
CC656933 963 bp DNA linear GSS 19-JUN-2003
LOCUS OGMDQ20TM ZM 0.7.1.5 KB Zea mays genomic clone ZMMBMA0554D15,
DEFINITION genomic survey sequence.
ACCESSION CC656933
VERSION CC656933.1 GI:32060225
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 963)
Whiteaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other_GSSs: OGMDQ20TV
Contact: Cathy Whiteaw
TIGR

FEATURES
source

9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteaw@tigr.org
Seq primer: TF
Class: sheared ends.
Location/Qualifiers
1..963
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0554D15"
/clone_lib="ZM 0.7.1.5 KB"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match 100.0%; Score 40; DB 9; Length 963;
Best Local Similarity 100.0%; Pred. No. 3.4e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCTCCAAACATCCATCTTACTCATGCAAC 40
|||||
836 AGGATACCTACTCTCCAAACATCCATCTTACTCATGCAAC 875
|||||

RESULT 4
H24137 528 bp mRNA linear EST 06-JUL-1995
LOCUS YMS0908.r1 Soares infant brain INIB Homo sapiens cDNA clone
DEFINITION IMAGE:51939 5', mRNA sequence.
ACCESSION H24137
VERSION H24137.1 GI:892832
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 528)
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevisan,E., Waterson,R., Williamson,A., Wohlmann,P. and
Wilson,R.
The WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson RK
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
Insert Size: 1161
High quality sequence stops: 349
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 1161 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 349.
Location/Qualifiers
1..528
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:424875"
/db_xref="taxon:9606"
/clone="IMAGE:51939"
/sex="female"
/dev_stage="73 days post natal"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares infant brain INIB"
/note="Organ: whole brain; Vector: Latmid B4; Site_1: Not

1; site 2; Hind III; 1st strand cDNA was primed with a Not I - oligo(dT) primer 15',
 AACTGAGAGATTCCGGCGCGCAATTTTCTTTTCTTTT 3');
 double-stranded cDNA was ligated to Hind III adaptor
 (Pharmacia), digested with Not I and directionally cloned
 into the Not I and Hind III sites of the lacZid BA vector.
 Library went through one round of normalization. Library
 constructed by Benito Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 59.5%; Score 23.8; DB 7; Length 528;
 Best Local Similarity 80.0%; Pred. No. 97;
 Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 ATACCTACTCCCAACATTCATCTTACTGATGCA 38
 |||||
 346 AATCTACTCTCAACATTCATCTTATACATGCA 380

RESULT 5

BZ706605/c

BZ706605 375 bp DNA linear GSS 18-MAY-2003
 SM416B1-G07_55_13.ab1 Spider Monkey genomic BAC library Ateles
 Geoffroyi genomic, genomic survey sequence.

ACCESSION BZ706605
 VERSION BZ706605
 KEYWORDS GSS.
 BZ706605.1 GI:30843142

SOURCE Ateles geoffroyi (black-handed spider monkey)
 ORGANISM Ateles geoffroyi

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Atelinae;
 Ateles.

REFERENCE 1 (bases 1 to 375)
 AUTHORS Qian, Y., Jin, L., and Su, B.
 TITLE BAC end sequences of spider monkey genomic library
 JOURNAL Unpublished (2003)
 COMMENT Contact: Qian Y, Jin L, Su B.
 Center for Genome Information
 University of Cincinnati
 Kettering Lab, 3223 Eden Ave., Cincinnati, OH 45267-0056, USA
 Tel: 1-513-558-6678
 Fax: 1-513-558-4505
 Email: subnucmail.uc.edu
 Seq primer: T7
 Class: BAC ends.

FEATURES

Location/Qualifiers
 1..375
 /organism="Ateles geoffroyi"
 /mol_type="genomic DNA"
 /strain="black-handed spider monkey"
 /db_xref="taxon:9509"
 /sex="Male"
 /issue_type="Skin"
 /cell_type="Fibroblast"
 /cell_line="AG05352"
 /dev_stage="3 DA"
 /clone_lib="Spider Monkey genomic BAC library"
 /note="Vector: pBACe3.6; Site 1: EcoRI; Genomic DNA was
 partially digested with EcoRI; Vector: pBACe3.6,
 Recombinants were transformed into DH10B."

ORIGIN
 Query Match 59.0%; Score 23.6; DB 8; Length 375;
 Best Local Similarity 74.4%; Pred. No. 1;e+02;
 Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

ORIGIN

QY 2 GGATACCTACTCCCAACATTCATCTTACTGATGCA 40
 |||||
 255 GGATATCTACTCTGTGAACCATTCATTTTCTNAGCAAC 217

Db 255 GGATATCTACTCTGTGAACCATTCATTTTCTNAGCAAC 217

RESULT 6

COL17589

LOCUS COL17589 738 bp mRNA linear EST 16-JUN-2004
 DEFINITION GR_EB01H15.r GR_Eb Gossypium raimondii cDNA clone GR_EB01H15 3',
 mRNA sequence.

ACCESSION COL17589
 VERSION COL17589.1 GI:48816276
 KEYWORDS EST.

SOURCE Gossypium raimondii
 ORGANISM Gossypium raimondii

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.

REFERENCE 1 (bases 1 to 738)
 AUTHORS Kim, H., Yu, Y., Kudrna, D., Hatfield, J., Stum, D., Mueller, C.,
 Udall, V.A., Kapp, R.A., Wendel, J.F., Rao, K., Soderlund, C. and
 Wing, R.A.

TITLE Global assembly of Cotton ESTs
 JOURNAL Unpublished (2004)
 COMMENT Contact: Rod A. Wing
 Arizona Genomics Institute
 The University of Arizona
 Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
 Tel: 520 626 9595
 Fax: 520 621 1259
 Email: http://genome.arizona.edu
 Plate: 01 row: H column: 15.

FEATURES

source

Location/Qualifiers
 1..738
 /organism="Gossypium raimondii"
 /mol_type="mRNA"
 /db_xref="taxon:29730"
 /clone="GR_EB01H15"
 /issue="GR_floral"
 /dev_stage="3 to +3 DPA"
 /lab_host="GR_ED"
 /clone_lib="GR_ED"
 /note="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2:
 EcoRV; Library made by invitrogen with RNA supplied by
 Wendle lab. Directional cloned into NotI-EV. Clonies
 plated/picked by AGI. More glycerol clones held in -80."

ORIGIN

Query Match 58.0%; Score 23.2; DB 7; Length 738;
 Best Local Similarity 77.8%; Pred. No. 1.8e+02;
 Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 ATACCTACTCCCAACATTCATCTTACTGATGCA 39
 |||||
 224 ATACACACTCCCATCAACGCAAGTACATGCA 259

RESULT 7

CRI181951

LOCUS CRI181951 783 bp DNA linear GSS 06-JUL-2004
 DEFINITION Reverse strand read from insert in 5'HPT insertion targeting and
 chromosome engineering clone MHPN61p21, genomic survey sequence.

ACCESSION CRI181951.1 GI:49960800
 VERSION CRI181951.1
 KEYWORDS GSS; genome survey sequence; MICEP.
 SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 783)
 AUTHORS Adams, D.J., Biggs, P.J., Cox, A.V., Davies, R.M., van der Weyden, L.,
 Jonkers, J., Smith, J., Plumb, R.W., Taylor, R.G., Nishijima, I., Yu, Y.,
 Rogers, J. and Bradley, A.

TITLE Direct Submission
 JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. http://www.sanger.ac.uk/MICEP
 location/Qualifiers
 1..783
 /organism="Mus musculus"
 /mol_type="genomic DNA"

```

/db_xref="taxon:10090"
/clone="MHPN361p21"
/clone_1ib="MHPN"

ORIGIN

Query Match      58.0%; Score 23.2; DB 9; Length 783;
Best Local Similarity 77.8%; Pred. No. 1.8e+02;
Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY
4 ATACCTACTCCCAACATCATCTTATCATGCAAC 39
|||||
514 ATACCTACTCTTAAGCAATGCATCTTACTCCAGAA 549

RESULT 8
CL109241/c      1018 bp      DNA      linear      GSS 05-JAN-2004
LOCUS
DEFINITION
ISB1-51D14 Sp6.1 ISB1 Xenopus tropicalis genomic clone ISB1-51D14,
genomic survey sequence.
ACCESSION
CL109241
VERSION
CL109241.1 GI:40602876
KEYWORDS
GSS.
SOURCE
Xenopus tropicalis (western clawed frog)
ORGANISM
Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 1018)
Kremetzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
Mardis, E. and Wilson, R.
A physical map of the xenopus tropicalis genome
Unpublished (2003)
COMMENT
Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: gsubmissions@watson.wustl.edu
Insert Length: 75000 Std Error: 0.00
Seq primer: Sp6 ATTTAGGTGACACTATAG
Class: BAC ends
High quality sequence start: 72
High quality sequence stop: 524.
Location/Qualifiers
1. 1018
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/db_xref="taxon:8364"
/clone="ISB1-51D14"
/clone_1ib="ISB1"
/note="Vector: pBelobAC11, ISB-1 Xenopus tropicalis BAC
Library Segment 1"

ORIGIN

Query Match      58.0%; Score 23.2; DB 9; Length 1018;
Best Local Similarity 77.8%; Pred. No. 1.9e+02;
Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY
5 TACCTACTCCCAACATCATCTTATCATGCAAC 40
|||||
156 TACATCTCCCAACATCATCTTATGAAATAC 121

RESULT 9
CD495749      1253 bp      mRNA      linear      EST 12-JUN-2003
LOCUS
DEFINITION
CD418-D05.y1d-8 SHGC-CDA Gasterosteus aculeatus cDNA clone
CD495749
CD495749.1 GI:31422780
KEYWORDS
EST.
SOURCE
Gasterosteus aculeatus (three spined stickleback)
Gasterosteus aculeatus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
```

```

Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes;
Gasterosteidae; Gasterosteus.
1 (bases 1 to 1253)
Kingsley, D.M., Peichel, C., Balabhadra, S., Grimwood, J., Dickson, M.,
Schmutz, J. and Myers, R.M.
Expressed sequence tags from Gasterosteus aculeatus
Unpublished (2003)
COMMENT
Contact: Kingsley, DM
HMI and Department of Developmental Biology
Stanford University School of Medicine
Beckman Center B300, 279 Campus Drive, Stanford, CA 94305-5329, USA
Tel: 650 725 5954
Fax: 650 725 7739
Email: kingsley@cmgm.stanford.edu
Plate: 18
High quality sequence start: 15
High quality sequence stop: 667.
Location/Qualifiers
1. 1253
/organism="Gasterosteus aculeatus"
/mol_type="mRNA"
/strain="Salinas river, CA"
/db_xref="Salinas:69293"
/clone="CD418-D05"
/sex="mixed male and female"
/tissue_type="heads and internal organs combined"
/dev_stage="adult"
/clone_1ib="SHGC-CDA"
/note="Vector: lambda ZAP Express/pBK-CMV, site_1: EcoRI
(5' adaptor); Site 2: XhoI (3' linker primer); The mixed
organ cDNA library was generated using the ZAP-CDNA method
by Stratagene. First strand cDNA synthesis was primed with
a 50 bp linker primer containing an oligo dT sequence
preceded by a synthetic XhoI site. 5 prime adaptors were
used containing an EcoRI cohesive end. The finished cDNAs
were inserted in to the ZAP express vector
unidirectionally in the sense orientation with respect to
the lacZ promoter of pBK-CMV. An amplified library was
prepared from approximately 3 million primary clones in
the lambda ZAP Express vector. In vivo excision was then
used to generate individual pBK-CMV phagemid clones for
EST sequencing."

ORIGIN

Query Match      58.0%; Score 23.2; DB 6; Length 1253;
Best Local Similarity 77.8%; Pred. No. 2e+02;
Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY
5 TACCTACTCCCAACATCATCTTACTCATGCAAC 40
|||||
1047 TACCTCTTCCCAACATCATCTTACCATCCAAC 1082

RESULT 10
AZ883408/c      405 bp      DNA      linear      GSS 05-MAR-2001
LOCUS
DEFINITION
RPCT-23-189L2.TV RPCT-23 Mus musculus genomic clone RPCT-23-189L2,
genomic survey sequence.
ACCESSION
AZ883408
AZ883408.1 GI:13202353
KEYWORDS
GSS.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 405)
Zhao, S., Nierman, W., Feldblum, T., Malek, J., Shatman, S.,
Akinre, B., Levins, M., McGinn, S., Teegay, G., Geer, K., Krol, M., de
Jong, P. and Fraser, C.M.
Mouse BAC End Sequences from Library RPCT-23
Unpublished (1999)
JOURNAL
COMMENT
Other GSSs: RPCT-23-189L2.TV
Contact: Shaying Zhao
```

Department of Eukaryotic Genomics
The Institute for Genomic Research
7712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: azha@tigr.org
Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources (<http://www.choi.org/bacpac/orderingframe.htm>). BAC end page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
Plate: 189 row: L column: 2
Seq primer: SP6
Class: BAC ends.

FEATURES

source
Location/Qualifiers

1..405
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-189L2"
/sex="Female"
/lab_host="DH10B"
/clone_1lb="RPCI-23"
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACe3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."

ORIGIN

Query Match 57.5%; Score 23; DB 8; Length 405;
Best Local Similarity 74.4%; Pred. No. 1.9e+02;
Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 AGGATACCTTCTCCCAACATCCATCTTATCATGCAA 39
DB 182 AGGTACCAAAATCCCAACATGCTTCTTATCATGCAA 144

RESULT 11
BG506830/c 798 bp mRNA linear EST 27-MAR-2001
LOCUS 601861443F1 NIH_MGC_77 Homo sapiens cDNA IMAGE:4071154 5',
DEFINITION mRNA sequence.
ACCESSION BG506830
VERSION BG506830.1 GI:13468347
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 798)
NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLM916 row: e column: 11
High quality sequence stop: 2.
Location/Qualifiers

FEATURES

source
Location/Qualifiers

1..798
/organism="Homo sapiens"
/mol_type="mRNA"

/db_xref="taxon:9606"
/clone="IMAGE:4071154"
/lab_host="DH10B (T1 phage-resistant)"
/clone_1lb="NIH_MGC_77"
/note="Organ: lung; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggcgatcgccg); Site 2: SfiI (ggcgatcgccg); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCGCATTTATGCGC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGCGCGCGCATATG-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.9 kb (range 0.5-4.0 kb). 12/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH-MGC Library."

ORIGIN

Query Match 57.5%; Score 23; DB 4; Length 798;
Best Local Similarity 83.9%; Pred. No. 2.2e+02;
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 10 ACTCCCAACATCCATCTTACTCATGCAAC 40
DB 222 ACTCCCAACAGCCTTCTTATGATGATC 192

RESULT 12
BF143726 968 bp mRNA linear EST 24-OCT-2000
LOCUS 601789606F1 NCI_CGAP_Lu30 Mus musculus cDNA clone IMAGE:4020600 5',
DEFINITION mRNA sequence.
ACCESSION BF143726
VERSION BF143726.1 GI:10982766
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclirognathi; Muridae; Murinae; Mus.
1 (bases 1 to 968)
NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLM9274 row: k column: 01
High quality sequence start: 3
High quality sequence stop: 648.
Location/Qualifiers

FEATURES

source

1..968
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CZECH 11"
/db_xref="taxon:10090"
/clone="IMAGE:4020600"
/tissue_type="tumor, metastatic to mammary"
/lab_host="DH10B"
/clone_1lb="NCI CGAP Lu30"
/note="Organ: lung; Vector: PCMV-SPORE; Site 1: NotI; Site 2: SalI; transgenic model WNT-1, expression driven by MMTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"

ORIGIN

Query Match 57.5%; Score 23; DB 2; Length 968;
Best Local Similarity 74.4%; Pred. No. 2.3e+02;
Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 2 GGATACCTACTCCCAACAATCATCTTACTATCGAAC 40
 |||||
 DB 818 GGCTCCGTACTCCCAACAACCAACCCACACACCCCAAC 856

RESULT 13
 BM406466
 LOCUS
 DEFINITION EST580781 potato roots Solanum tuberosum cDNA clone cPROZ7F14 5'
 end, mRNA sequence.
 ACCESSION BM406466
 VERSION BM406466.1 GI:18258084
 KEYWORDS EST.
 SOURCE Solanum tuberosum (potato)
 ORGANISM Solanum tuberosum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanales; Solanales; Solanum.
 1 (bases 1 to 771)
 van der Hoeven, R., Sun, H., Karamycheva, S.A., Tsai, J., Van Aken, S., Utechtback, T., Chiemingo, A., Bougri, O., Buell, C.R., Romning, C., Tanksley, S. and Baker, B.
 Generation of ESTs from potato roots
 Unpublished (2001)
 Contact: Robin Buell
 The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-array@tigr.org
 This clone can be obtained from the University of Arizona Genomics Institute. Orders can be made through URL:
<http://genome.arizona.edu/orders/>
 Seq primer: T3.

FEATURES
 source
 Location/Qualifiers
 1..771
 /organism="Solanum tuberosum"
 /mol_type="mRNA"
 /cultivar="Kennebec"
 /db_xref="taxon:4113"
 /clone="cPROZ7F14"
 /tissue_type="roots"
 /dev_stage="in vitro grown stem cuttings"
 /lab_host="SOLR"
 /clone_lib="potato roots"
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; supplier: Cornell University, Tanksley lab. sequencing; The Institute for Genomic Research. Roots were isolated from in vitro grown stem cuttings on CM medium. Roots were isolated two weeks after placing the stem cuttings from in vitro grown plants on medium."

ORIGIN

Query Match 57.0%; Score 22.8; DB 4; Length 771;
 Best Local Similarity 79.4%; Pred. No. 2.6e+02;
 Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 5 TACCTACTCCCAACAATCATCTTACTATCGAAC 38
 |||||
 DB 649 TCCCTCTTCACACATTCACATCTCTCAAGCA 682

RESULT 14
 AM257178
 LOCUS
 DEFINITION EST305315 KV2 Medicago truncatula cDNA clone KV2-7D15, mRNA
 sequence.
 ACCESSION AM257178
 VERSION AM257178.1 GI:6605435
 KEYWORDS EST.
 SOURCE Medicago truncatula (barrel medic)
 ORGANISM Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eudicots I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae; Medicago.
 1 (bases 1 to 410)
 Vandenbosch, K., Hurt, J., Moore, J., Beremand, P., Peng, H., Ellis, L., Town, C.D., Bowman, C.L., Craven, M.B., Hansen, T.S., Holt, I.E. and Frazer, C.M.
 ESTs from roots of Medicago truncatula after Rhizobium inoculation
 Unpublished (1999)
 Contact: Vandenbosch K
 Department of Plant Biology
 University of Minnesota
 220 Biosci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA
 Tel: 612 624 2755
 Fax: 612 625 1738
 Email: kvandenbosch.umn.edu
 Texas A&M EST name: T115883e
 TIGR sequence name: MTAA0207K
 More information is available at: . (and for clone ordering info)
<http://chryslr.tamu.edu/medicago>
 Seq primer: SKmod (CTA GAA CTA gtc gAT CC).
 Seq primer: SKmod (CTA GAA CTA gtc gAT CC).

FEATURES
 source
 Location/Qualifiers
 1..410
 /organism="Medicago truncatula"
 /mol_type="mRNA"
 /cultivar="genotype A17"
 /db_xref="taxon:3880"
 /clone="KV2-7D15"
 /tissue_type="Seedling roots"
 /dev_stage="2 days post-inoculation with Sinorhizobium meliloti"
 /lab_host="E. coli strain SOLR"
 /clone_lib="KV2"
 /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into the Uniap XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-Zap phage using Ex-assist helper phage and propagated in SOLR cells."

ORIGIN

Query Match 56.5%; Score 22.6; DB 2; Length 410;
 Best Local Similarity 86.2%; Pred. No. 2.8e+02;
 Matches 25; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 ACCTACTCCCAACAATCATCTTACTCA 34
 |||||
 DB 16 ACCAATCTCCCAACAATCATCTGCTCA 44

RESULT 15
 BE205271
 LOCUS
 DEFINITION EST397947 KV0 Medicago truncatula cDNA clone KV0-21123, mRNA
 sequence.
 ACCESSION BE205271
 VERSION BE205271.1 GI:8748567
 KEYWORDS EST.
 SOURCE Medicago truncatula (barrel medic)
 ORGANISM Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eudicots I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae; Medicago.
 1 (bases 1 to 560)
 Vandenbosch, K., Endre, G., Hur, J., Moore, J., Beremand, P., Ellis, L., Town, C.D., Bowman, C.L., Craven, M.B., Hansen, T.S., Holt, I.E. and Frazer, C.M.
 ESTs from uninoculated seedling roots of Medicago truncatula
 Unpublished (1999)
 Contact: Vandenbosch K
 Department of Plant Biology
 University of Minnesota

220 Biosci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA
 Tel: 612 624 2755
 Fax: 612 625 1738

Email: kvandenb@cbs.umn.edu
 Texas A&M University name: T265043e
 TIGR sequence name: MTGBK60TK

More information is available at...

<http://chryslr.tamu.edu/medicago>
 Seq primer: Skmod (CTA gaa cta gtc gat cc).

FEATURES

source

```

1..560
    location/Qualifiers
    /organism="Medicago truncatula"
    /mol_type="mRNA"
    /cultivar="genotype A17"
    /db_xref="taxon:3880"
    /clone="pKV0-21123"
    /issue_type="Seedling roots"
    /dev_stage="Immediately prior to inoculation with
    Sinorhizobium meliloti (0 hour)"
    /lab_host="E.coli strain XL0LR"
    /clone_lib="KV0"
    /note="Vector: Bluescript SK-; Site_1: EcoRI; Site_2:
    XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
    was directionally ligated into the Unizap XR vector from
    Stratagene and packaged using Gigapack III Gold packaging
    extracts. Plasmids containing cDNA inserts were excised
    from the recombinant lambda-Zap phage using Ex-assist
    helper phage and propagated in XL0LR cells."
  
```

ORIGIN

Query Match

Best Local Similarity 86.5%; Score 22.6; DB 2; Length 560;
 Matches 25; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 ACCTACTCCCAACATCCATCTTACTCA 34
 ||| ||| ||| ||| ||| ||| ||| |||
 DB 24 ACCAATCCCAACATCCATCTGCTCA 52

Search completed: September 16, 2005, 08:08:43
 Job time : 124.754 secs

This page blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd

OM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 08:25:31 ; Search time 4399.14 Seconds

14440.280 Million cell updates/sec

Title: US-10-713-381-2_COPY_1_1311
Page: 1311

Sequence: 1 ccattggtgtctctatgaaaa.....cgcacatatgtcttcctgaac 1311

Scoring table: IDENTITY_NUC

Searched: 4708233 beqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

```
Minimum DB seq length: 0
```

3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842

Post-processing:	Minimum Match	0%
	Maximum Match	100%

Listing first 45 summaries

Database : GenEmbl : *

```

1:  gb_da: *
2:  gb_htg: *
3:  gb_in: *
4:  gb_om: *
5:  gb_ov: *
6:  gb_pat: *
7:  gb_ph: *
8:  gb_pl: *
9:  gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_by: *
13: gb_un: *
14: gb_vl: *

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1311	100.0	1394	6	AX224395	AX224395 Sequence
2	1311	100.0	1394	6	AX224395	AX224395 Sequence
3	1311	100.0	1394	6	BD062176	BD062176 Male f18
4	1311	100.0	1394	6	BD062177	BD062177 Male f18
5	1311	100.0	3343	8	AF360356	AF360356 Zea mays
6	150.6	12.3	255	6	AX224392	AX224392 Sequence
7	146	11.1	158	6	AX224396	AX224396 Sequence
8	77	5.9	186159	2	AC147602	AC147602 Zea mays
9	73.4	5.6	25769	2	AC117267	AC117267 Dictyostel
10	71.2	5.4	713	3	AF034389	AF034389 Plasmidori
11	71.2	5.4	3576	3	DD086962	DD086962 Plasmidori
12	70.8	5.4	145526	2	EX957346	UB6962 Dictyostel
13	70.4	5.4	347582	3	PFMAL4P1	BA8957346 Danclo re
14	70.2	5.4	883	6	CQ422641	AL034557 Plasmodii
15	70	5.3	241	11	BV119878	CQ422641 Sequence
16	70	5.3	247	11	BV119874	BV119878 Sequence
17	70	5.3	248	11	BV119877	BV119874 PZA0137
18	69.8	5.3	13400	2	C0855052	BV119877 PZA0137
19	69.6	5.3	13583	2	AC151050	C0855052 Sequence
						AC151050 Zea mays

C	20	69.4	5.3	110000	8	CR382132	17	Continuation (18 of 19)
C	21	69.4	5.3	181381	2	CR450818	Danto rer	
C	22	69.2	5.3	114817	2	AC149936	AC1499836 Zea mays	
C	23	69.2	5.3	266544	3	AC116956	AC1169596 Dicotyleste	
C	24	69	5.3	164443	2	CR751221	Danto rer	
C	25	68.8	5.2	7347	1	AF211124	AF211124 Carone11	
C	26	68.6	5.2	186199	2	AC147602	AC147602 Zea mays	
C	27	68.4	5.2	213	11	BV119875	BV119875 PZAO1377	
C	28	68.4	5.2	216	11	BV119871	BV119871 PZAO1377	
C	29	68.4	5.2	224	11	BV119882	BV119882 PZAO1377	
C	30	68.4	5.2	245	11	BV119881	BV119881 PZAO1377	
C	31	68.4	5.2	246	11	BV119869	BV119869 PZAO1377	
C	32	68.2	5.2	346296	8	AF090447	AF090447 Zea mays	
C	33	68	5.2	110000	2	PFMAL7P1	02 Continuation (3 of 4)	
C	34	67.8	5.2	227495	2	AC148480	AC148480 Zea mays	
C	35	67.6	5.2	351	11	BV110530	BV110530 PZAO1059	
C	36	67.6	5.2	331039	3	AC116979	AC116979 Dicotyleste	
C	37	67.4	5.1	170	11	BV119873	BV119873 PZAO1377	
C	38	67	5.1	2781	3	AY044085	AY044085 Dicotyleste	
C	39	67	5.1	110000	3	AC116984	1 Continuation (2 of 3)	
C	40	66.8	5.1	228	11	BV119870	BV119870 PZAO1377	
C	41	66.8	5.1	234	11	BV119883	BV119883 PZAO1377	
C	42	66.8	5.1	7774	12	AF336825	AF336825 Cloning	
C	43	66.8	5.1	8658	12	AF336824	AF336824 Cloning	
C	44	66.8	5.1	12029	3	AE001434	AE001434 Plasmodiu	
C	45	66.8	5.1	47573	3	AF030694	AF030694 Plasmodiu	

ALIGNMENTS

[illegible]

Db 181 CTGACCGAGGCCCATGAGACACCTTTGCGGAGACCCATCAAGGCGCTTTCGATGGCCCA 240
Qy 241 CGAGAGCTATCGGGTGTGGTGAATCCAGGGAGATATATGCCCCAGAAATGCTACTATA 300
Db 241 CGAGAGCTATCGGGTGTGGTGAATCCAGGGAGATATATGCCCCAGAAATGCTACTATA 300
Qy 301 TTATATTTCTTTAGATATTTATTTATTTTGGAAAAATAACAATTAATCTTTTGTGA 360
Db 301 TTATATTTCTTTAGATATTTATTTTGGAAAAATAACAATTAATCTTTTGTGA 360
Qy 361 GGGGCTTCAGATAGATTTTCGCTTAAAGGCCAGAAATGCGAGAGCAAGCCATGCTAAGT 420
Db 361 GGGGCTTCAGATAGATTTTCGCTTAAAGGCCAGAAATGCGAGAGCAAGCCATGCTAAGT 420
Qy 421 TCCACTATTGGGAGCTACCGAGAACAGATTTAAATAAATGCAAGTAATCTAATCCACT 480
Db 421 TCCACTATTGGGAGCTACCGAGAACAGATTTAAATAAATGCAAGTAATCTAATCCACT 480
Qy 481 CGAAGCTATCATGTATGTTTAAAGAAACATCTATTTAAACACAGATCCTCTTAAAAA 540
Db 481 CGAAGCTATCATGTATGTTTAAAGAAACATCTATTTAAACACAGATCCTCTTAAAAA 540
Qy 541 CAAGCATATTTGAAAGAGACAAATTAATGTTACAGTTTACAAACATCTAAGCGACAAA 600
Db 541 CAAGCATATTTGAAAGAGACAAATTAATGTTACAGTTTACAAACATCTAAGCGACAAA 600
Qy 601 TTATATCGAAAGGTAAAGCTATGAGTTCAGATTTTCTTTTCAATCTGTTATTTGTT 660
Db 601 TTATATCGAAAGGTAAAGCTATGAGTTCAGATTTTCTTTTCAATCTGTTATTTGTT 660
Qy 661 ATTGTTTTATATACATTTTCTTCTTACAAATGAGTGAATTTCTTCGATTTTAA 720
Db 661 ATTGTTTTATATACATTTTCTTCTTACAAATGAGTGAATTTCTTCGATTTTAA 720
Qy 721 ATGACATATAAGCTATTTTATATAAGAGACGCAATGCTGATCTGCTTCAAAAATC 780
Db 721 ATGACATATAAGCTATTTTATATAAGAGACGCAATGCTGATCTGCTTCAAAAATC 780
Qy 781 TTTCTGATTTTAAAGCTAGTTGGCAACCTGTTTCTTCAAGAATTTGATTT 840
Db 781 TTTCTGATTTTAAAGCTAGTTGGCAACCTGTTTCTTCAAGAATTTGATTT 840
Qy 841 TTTCAAAAAAATAGTTTATTTTCTTTTAAATAAGAAAACTTAGAAAAATAGAGT 900
Db 841 TTTCAAAAAAATAGTTTATTTTCTTTTAAATAAGAAAACTTAGAAAAATAGAGT 900
Qy 901 TGGCAGACTAGCCCTAGAAATGTTTCCCAATTAATTAACAATCTGTGTATAATTTTG 960
Db 901 TGGCAGACTAGCCCTAGAAATGTTTCCCAATTAATTAACAATCTGTGTATAATTTTG 960
Qy 961 GCCAGCCCCATAAATATTAAACGAAACTGAATCGAGCAACCAATCTGAGCTAT 1020
Db 961 GCCAGCCCCATAAATATTAAACGAAACTGAATCGAGCAACCAATCTGAGCTAT 1020
Qy 1021 TTTCTAGATTAGTAAAG 1080
Db 1021 TTTCTAGATTAGTAAAG 1080
Qy 1081 TGGAGATGTGGGTTTGGCAACATAGCCATTAATCAATGCTAGTGTGCTTACGTA 1140
Db 1081 TGGAGATGTGGGTTTGGCAACATAGCCATTAATCAATGCTAGTGTGCTTACGTA 1140
Qy 1141 GGGTGGGAGCTGTGCTGATCTGATGAGATCTAATGCTTGTGAACGGTTCGTC 1200
Db 1141 GGGTGGGAGCTGTGCTGATCTGATGAGATCTAATGCTTGTGAACGGTTCGTC 1200
Qy 1201 TTGTTTCATCTGTCAGAGCCCTGCTATTTGGAACCAAGAGATACTATCCCAAAACAT 1260
Db 1201 TTGTTTCATCTGTCAGAGCCCTGCTATTTGGAACCAAGAGATACTATCCCAAAACAT 1260
Qy 1261 CCATCTTACTCATGCAACTTCCATGCAACACGACATATGTTTCTGTAAC 1311
Db 1261 CCATCTTACTCATGCAACTTCCATGCAACACGACATATGTTTCTGTAAC 1311

RESULT 2
AX224395
LOCUS AX224395 1394 bp DNA linear PAT 10-SEP-2001
DEFINITION Sequence 2 from Patent WO0160997.
ACCESSION AX224395
VERSION AX224395.1 GI:15554637
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1
AUTHORS Albertsen, M.C., Fox, T.W., Garnaat, C.W., Huffman, G. and Kendall, T.L.
TITLE Male tissue-preferred regulatory region and method of using same
JOURNAL Patent: WO 0160997-A 2 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
source location/Qualifiers
1..1394
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"
ORIGIN
Query Match 100.0%; Score 1311; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 3.6e-252;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CCATGTCCTCTATGAAAAAGATAGTACATATGCTATATCCGTTTCTTAAAGGTC 60
Db 1 CCATGTCCTCTATGAAAAAGATAGTACATATGCTATATCCGTTTCTTAAAGGTC 60
Qy 61 CTTCCTTGCCCTTATTAATGATGATGCGGGTTACAAAAAATTCCACGCGTGCATAT 120
Db 61 CTTCCTTGCCCTTATTAATGATGATGCGGGTTACAAAAAATTCCACGCGTGCATAT 120
Qy 121 CTCGATGTTCCACTCTCCACCTCGCGTTGACATTTCTTGGAGTGGTGGTCCAT 180
Db 121 CTCGATGTTCCACTCTCCACCTCGCGTTGACATTTCTTGGAGTGGTGGTCCAT 180
Qy 181 CTGACCGAGGCCCATACAGACACTTTCCGGACACCCATCAAGGCGCTTCCGATGCGCA 240
Db 181 CTGACCGAGGCCCATACAGACACTTTCCGGACACCCATCAAGGCGCTTCCGATGCGCA 240
Qy 241 CGAGCGATAGGGGTGCTGTGATCCAGGGATATATGTCCTCCCAATTCGTACCTATA 300
Db 241 CGAGCGATAGGGGTGCTGTGATCCAGGGATATATGTCCTCCCAATTCGTACCTATA 300
Qy 301 TTATTAATCTTATGATATTTATTTTGGAAAAATTAACAATTAATTAATCTTTGTGTA 360
Db 301 TTATTAATCTTATGATATTTATTTTGGAAAAATTAACAATTAATTAATCTTTGTGTA 360
Qy 361 GGGGCTTCAGATAGATTTTTCGTTAGGCCCCAGAAATCGAGAGACAGCATGCTAAGT 420
Db 361 GGGGCTTCAGATAGATTTTTCGTTAGGCCCCAGAAATCGAGAGACAGCATGCTAAGT 420
Qy 421 TCCACTATTGGGAGCTACCGAGAACAGATTTAAATAAATGCAAGTAATCTAATCCACT 480
Db 421 TCCACTATTGGGAGCTACCGAGAACAGATTTAAATAAATGCAAGTAATCTAATCCACT 480
Qy 481 CGAAGCTATCATGTATGTTTAAAGAAACATCTATTTAAACACAGATCCTCTTAAAAA 540
Db 481 CGAAGCTATCATGTATGTTTAAAGAAACATCTATTTAAACACAGATCCTCTTAAAAA 540
Qy 541 CAAGCATATTTGAAAGAGACAAATTAATGTTACAGTTTACAAACATCTAAGGCGACAAA 600
Db 541 CAAGCATATTTGAAAGAGACAAATTAATGTTACAGTTTACAAACATCTAAGGCGACAAA 600
Qy 601 TTATATCGAAAGGTAAAGCTATGAGTTCAGATTTTCTTTTCAATCTGTTATTTGTT 660
Db 601 TTATATCGAAAGGTAAAGCTATGAGTTCAGATTTTCTTTTCAATCTGTTATTTGTT 660

QY 661 ATTGTTTATATACATTTTCTCTCTTACATAGATGATTTTCTCCGATTTTATAAA 720
 DB 661 ATTTGTTTATATACATTTTCTCTCTTACATAGATGATTTTCTCCGATTTTATAAA 720
 QY 721 ATGACTATAAGTCATTTTATATAAGAGACGCGATGCTGATGATTCCTGTTCAAAAATC 780
 DB 721 ATGACTATAAGTCATTTTATATAAGAGACGCGATGCTGATGATTCCTGTTCAAAAATC 780
 QY 781 TTTCTGATTTTATAGAGCTAGTTGGCAACCCCTGTTCTTCAAGAAATTTGATTTT 840
 DB 781 TTTCTGATTTTATAGAGCTAGTTGGCAACCCCTGTTCTTCAAGAAATTTGATTTT 840
 QY 841 TTCAAAAAAATAGTTTATTTCTCTTATATAAATGAAAAACCTTAGAAAAATAGAT 900
 DB 841 TTCAAAAAAATAGTTTATTTCTCTTATATAAATGAAAAACCTTAGAAAAATAGAT 900
 QY 901 TGCCAGACTAGCCCTAGAAATGTTTCCCAATTAATTCATCACTGTGATTAATTTTG 960
 DB 901 TGCCAGACTAGCCCTAGAAATGTTTCCCAATTAATTCATCACTGTGATTAATTTTG 960
 QY 961 GCCAGCCCATTAATTTATTAACCGAATCTGAATCGAGCAAACTGAGCTAT 1020
 DB 961 GCCAGCCCATTAATTTATTAACCGAATCTGAATCGAGCAAACTGAGCTAT 1020
 QY 1021 TTCTCTAGATTTAGTAAAGAGAGAGAGAGAGAAATCACTTTTAATGATTCCTCC 1080
 DB 1021 TTCTCTAGATTTAGTAAAGAGAGAGAGAGAGAAATCACTTTTAATGATTCCTCC 1080
 QY 1081 TGAGATGTGCGGTTTGGAAGAGATAGCCACCGTATCATAGTGTGCTTACGTCGA 1140
 DB 1081 TGAGATGTGCGGTTTGGAAGAGATAGCCACCGTATCATAGTGTGCTTACGTCGA 1140
 QY 1141 GGTTCGGACGCTCTGTCATCTCACATGCACTACTGCTTGTTCACCGTTCGTC 1200
 DB 1141 GGTTCGGACGCTCTGTCATCTCACATGCACTACTGCTTGTTCACCGTTCGTC 1200
 QY 1201 TTGTTCCATGCTGTCAGAGCTTGGCTATTCGAAACCAAGAGATACCTTCCCAACAAT 1260
 DB 1201 TTGTTCCATGCTGTCAGAGCTTGGCTATTCGAAACCAAGAGATACCTTCCCAACAAT 1260
 QY 1261 CCATCTTACTCATGCACTTCATGCAACCGACATATGTTTCTCGAAC 1311
 DB 1261 CCATCTTACTCATGCACTTCATGCAACCGACATATGTTTCTCGAAC 1311
 RESULT 3
 LOCUS BD062176 1394 bp DNA linear PART 27-AUG-2002
 DEFINITION Male tissue-preferred regulatory region and method of using same.
 ACCESSION BD062176
 VERSION BD062176.1 GI:2260781
 KEYWORDS JP 2001520523-A/1.
 SOURCE Homo sapiens (human)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1394)
 AUTHORS Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G.A. and Kendall,T.L.
 TITLE Male tissue-preferred regulatory region and method of using same
 JOURNAL Patent: JP 2001520523-A 1 30-OCT-2001;
 COMMENT PIONEER HI BRED INTERNATIONAL INC
 PN JP 2001520523-A/1
 PD 30-OCT-2001
 PF 19-JUN-1998 JP 199504910
 PR 23-JUN-1997 US 08/880499
 PI MARC C ALBERTSEN,TIMOTHY W FOX,CARL W GARNAAAT,GARY A HUFFMAN,
 PI TIMMY L KENDALL
 PC C12N15/82,C12N15/29,C12N9/24,C12N9/22,C12N9/10,C12N9/00 PC
 PC A01H5/00
 CC Strandedness: Single;

CC Topology: Linear;
 FH Key Location/Qualifiers.
 FEATURES
 source Location/Qualifiers
 1..1394
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Query Match 100.0%; Score 1311; DB 6; Length 1394;
 Best Local Similarity 100.0%; Pred. No. 3,6e-252;
 Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CCATGCTGCTCTATGAAAAAGATGATACATGCTCTATATCCGTTTCTTAGAGGTC 60
 DB 1 CCATGCTGCTCTATGAAAAAGATGATACATGCTCTATATCCGTTTCTTAGAGGTC 60
 QY 61 CTTCCTGCTGCTTATGACTGAAATCGGGGTTTACAAAACTTCCAGGGTGCAATGAT 120
 DB 61 CTTCCTGCTGCTTATGACTGAAATCGGGGTTTACAAAACTTCCAGGGTGCAATGAT 120
 QY 121 CTCGATGTTCCACTTCTCCCACTCGCGTTGCACTTTCTTGAGTGTGCGTGGTCCAT 180
 DB 121 CTCGATGTTCCACTTCTCCCACTCGCGTTGCACTTTCTTGAGTGTGCGTGGTCCAT 180
 QY 181 CTGACCGAGGCCCATCAGACACTTTGCGGACACCCATCAAGGCTTTGCGATGGCCCA 240
 DB 181 CTGACCGAGGCCCATCAGACACTTTGCGGACACCCATCAAGGCTTTGCGATGGCCCA 240
 QY 241 CGAGAGGTATCGGGTGTGCTGATGCCAGGGGATATATGTCGCCCAATGTGCACTATA 300
 DB 241 CGAGAGGTATCGGGTGTGCTGATGCCAGGGGATATATGTCGCCCAATGTGCACTATA 300
 QY 301 TTATTAATCTTATGATATTTATTAATTTTGGAAAAATACAACTTAATACCTTTTGAT 360
 DB 301 TTATTAATCTTATGATATTTATTAATTTTGGAAAAATACAACTTAATACCTTTTGAT 360
 QY 361 GGGCTCAGCATGATTTTTCGCTTAGGGCCCGAAGATGCGAGACCAAGCATGTCTAGTG 420
 DB 361 GGGCTCAGCATGATTTTTCGCTTAGGGCCCGAAGATGCGAGACCAAGCATGTCTAGTG 420
 QY 421 TCCACTATTTGGGACTCCAGAACAGATTTTAAAAAATATACCAAGTAATCAATCCACT 480
 DB 421 TCCACTATTTGGGACTCCAGAACAGATTTTAAAAAATATACCAAGTAATCAATCCACT 480
 QY 481 CGAAGCTATCATGTATGTTTAAAGAAACATCTATTAAGACGATCCTCTTAAGAAA 540
 DB 481 CGAAGCTATCATGTATGTTTAAAGAAACATCTATTAAGACGATCCTCTTAAGAAA 540
 QY 541 CAAGCATTTTGAAGAGACAAATATGTACAGTTTACAAACATCTAAGAGCGACAAA 600
 DB 541 CAAGCATTTTGAAGAGACAAATATGTACAGTTTACAAACATCTAAGAGCGACAAA 600
 QY 601 TTATATGAAAGGTATGATGAGTTCAGATTTTCTTTTCATCTGTTATTTGTT 660
 DB 601 TTATATGAAAGGTATGATGAGTTCAGATTTTCTTTTCATCTGTTATTTGTT 660
 QY 661 ATTTGTTTATATACATTTTCTTCTTACATAGAGATTTTCTTCGATTTTATAA 720
 DB 661 ATTTGTTTATATACATTTTCTTCTTACATAGAGATTTTCTTCGATTTTATAA 720
 QY 721 ATGACTATAAGTCATTTTATATAAGAGACGCGATGCTGATGATTCCTGTTCAAAAATC 780
 DB 721 ATGACTATAAGTCATTTTATATAAGAGACGCGATGCTGATGATTCCTGTTCAAAAATC 780
 QY 781 TTTCTGATTTTATAGAGCTAGTTGGCAACCCCTGTTCTTCAAGAAATTTGATTTT 840
 DB 781 TTTCTGATTTTATAGAGCTAGTTGGCAACCCCTGTTCTTCAAGAAATTTGATTTT 840
 QY 841 TTCAAAAAAATAGTTTATTTCTCTTATATAAATGAAAAACCTTAGAAAAATAGAT 900
 DB 841 TTCAAAAAAATAGTTTATTTCTCTTATATAAATGAAAAACCTTAGAAAAATAGAT 900

QY	901	TGCCAGACTAGCCCTTAGAATGTTTCCCAATTAATTACATCATCTGTATTAATTATTG	960
Db	901	TGCCAGACTAGCCCTTAGAATGTTTCCCAATTAATTACATCACGTATTAATTATTG	960
QY	961	GCCAGCCCCATTAATTATTAAACCGAATCGAANATGAGGAAACCAATCTAGCAT	1020
Db	961	GCCAGCCCCATTAATTATTAAACCGAATCGAANATGAGGAAACCAATCTAGCAT	1020
QY	1021	TTCTCTAGATTAGTAAAAAGGAGAGAGAGAAATCATGTTTAACTATGTCCTCC	1080
Db	1021	TTCTCTAATTAGTAAAAAGGAGAGAGAGAAATCATGTTTAACTATGTCCTCC	1080
QY	1081	TGAGATGTCGGCTTTGGCAAGATAGCCACCGTATCTATAGCTCATAGTCTCATGCTCA	1140
Db	1081	TGAGATGTCGGCTTTGGCAAGATAGCCACCGTATCTATAGCTCATAGTCTCATGCTCA	1140
QY	1141	GGTTGGGAGGCTCGTGCTCATCTCAATGGCATCTACATGCTGTTCACACGTTGGTC	1200
Db	1141	GGTTGGGAGGCTCTGCTCATCTCAATGGCATCTACATGCTGTTCACACGTTGGTC	1200
QY	1201	TTGTTTCATCTGTCGAAGCTTGCTATTCTGAACCAAGAGATACCTACTCCAAACAT	1260
Db	1201	TTGTTTCATCTGTCGAAGCTTGCTATTCTGAACCAAGAGATACCTACTCCAAACAT	1260
QY	1261	CCATTTTACTCATGCACTTCCATGCAACACGCAATATGTTCTCTGAAC	1311
Db	1261	CCATTTTACTCATGCACTTCCATGCAACACGCAATATGTTCTCTGAAC	1311

	RESULT 4			
	BD062177			
LOCUS	BD062177	1394 bp	DNA	linear PAT 27-AUG-2002
DEFINITION	Male tissue-preferred regulatory region and method of using same.			
ACCESSION	BD062177			
VERSION	BD062177.1	GI:22607782		
KEYWORDS	JP 2001520523-A/2.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
	1 (bases 1 to 1394)			
	Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G.A. and			
	Kendall,T.L.			
TITLE	Male tissue-preferred regulatory region and method of using same			
JOURNAL	Patent: JP 2001520523-A 2 30-OCT-2001;			
	PIONEER HI BRED INTERNATIONAL INC			
COMMENT	PN JP 2001520523-A/2			
	PD 30-OCT-2001			
	PF 19-JUN-1998 JP 1998504910			
	PR 23-JUN-1997 US 08/880499			
	PI MARC C ALBERTSEN,TIMOTHY W FOX,CARL W GARNAAT,GARY A HUFFMAN,			
	PI TIMMY L KENDALL			
	PC C12N15/22,C12N15/29,C12N9/24,C12N9/22,C12N9/10,C12N9/00 PC			
	,C07K4/34,C12Q1/68,			
	PC A01H5/00			
	CC Seriadness: Single;			
	CC Topology: linear;			
FEATURES	FH Key	Location/Qualifiers.		
source	1..1394	Location/Qualifiers		
	/organism="Homo sapiens"			
	/mol_type="genomic DNA"			
	/db_xref="taxon:9606"			
ORIGIN				

Query Match	100.0%	Score 1311	DB 6	Length 1394
Best Local Similarity	100.0%	Pred. No. 3.6e-252		
Matches 1311	Conservative 0	Mismatches 0	Indels 0	Gaps 0

Oy	1	CCATGGTCTCTCTATGAAAAGATGAGACACATGCTATATCCGTTTCTTAGGGTCC	60
Db	1	CCATGGTCTCTCTATGAAAAGATGAGACACATGCTATATCCGTTTCTTAGGGTCC	60

QY	61	CTTCTTCGCTTATTTACTGA	CTGAATCGGGGT	TACAAA	AAACCTTCCACGGGTGCATGAT	120
DB	61	CTTCTTCGCTTATTTACTGA	CTGAATCGGGGT	TACAAA	AAACCTTCCACGGGTGCATGAT	120
QY	121	CTCCATGTTCACTTCTCCCA	CTTCGGGTGACATTTCTTGGA	TGTCGGTGGTTCCCAT	180	
DB	121	CTCCATGTTCACTTCTCCCA	CTTCGGGTGACATTTCTTGGA	TGTCGGTGGTTCCCAT	180	
QY	181	CTGACCGAGGCCCATCAGAC	ACCTTTGGGAGACACCATTCAGAGGCGCTTTCGGATGGGCCA	240		
DB	181	CTGACCGAGGCCCATCAGAC	ACCTTTGGGAGACACCATTCAGAGGCGCTTTCGGATGGGCCA	240		
QY	241	CGAGACGTATCGGGTGTGTG	TATCCAGGGGATATATGTCCCCCA	CAATGTGTCACTATA	300	
DB	241	CGAGACGTATCGGGTGTGTG	TATCCAGGGGATATATGTCCCCCA	CAATGTGTCACTATA	300	
QY	301	TTATATCTTTAGATATTTAT	TTATTTTGGAAAAATPACAACTATACCTTTTGAT	360		
DB	301	TTATATCTTTAGATATTTAT	TTATTTTGGAAAAATPACAACTATACCTTTTGAT	360		
QY	361	GGGCGTCAGACATAGATTTT	CGCTTGGGCCCGCAGAAATCGAGAGACAGCGCATGTCTAGTG	420		
DB	361	GGGCGTCAGACATAGATTTT	CGCTTGGGCCCGCAGAAATCGAGAGACAGCGCATGTCTAGTG	420		
QY	421	TCCACTATTTGCGACTACCC	ACAGATTTTAAAAAATTAACCAAGTAACTAATTCACCT	480		
DB	421	TCCACTATTTGCGACTACCC	ACAGATTTTAAAAAATTAACCAAGTAACTAATTCACCT	480		
QY	481	CGAAAGCTATCATGTATGTA	TGTTTAAAGAAACATCTTTTAAAAACAGATCTCTTTAAAAA	540		
DB	481	CGAAAGCTATCATGTATGTA	TGTTTAAAGAAACATCTTTTAAAAACAGATCTCTTTAAAAA	540		
QY	541	CAAGCATATTTCAAAAGAC	CAAAATATATGTTTACAGTTTACAAACATCTPAGAGCACA	600		
DB	541	CAAGCATATTTCAAAAGAC	CAAAATATATGTTTACAGTTTACAAACATCTPAGAGCACA	600		
QY	601	TTATATCGAAAGTAAAGCT	ATGACGTTTCCTTTTCATTCCTTGATTTTGT	660		
DB	601	TTATATCGAAAGTAAAGCT	ATGACGTTTCCTTTTCATTCCTTGATTTTGT	660		
QY	661	ATTGTTTTTATATACATTTT	CTCTCTTACAAATAGAGGATTTTCTTCGATTTTATAA	720		
DB	661	ATTGTTTTTATATACATTTT	CTCTCTTACAAATAGAGGATTTTCTTCGATTTTATAA	720		
QY	721	ATGACATAAAGTCAATTTT	TATATPAGAGCAGCATGTCTGTGATCTTCGTTCAAAATC	780		
DB	721	ATGACATAAAGTCAATTTT	TATATPAGAGCAGCATGTCTGTGATCTTCGTTCAAAATC	780		
QY	781	TTTCTGATTTTTTTAAGCT	ATGTTGGCAACCTGTTCCTTCAAGAAATTTGATTT	840		
DB	781	TTTCTGATTTTTTTAAGCT	ATGTTGGCAACCTGTTCCTTCAAGAAATTTGATTT	840		
QY	841	TTTCAAAAAAATTAAGTTAT	TTTTCTCTTATTAATAATAGAAAAACATTTAGAAAAAATAGAT	900		
DB	841	TTTCAAAAAAATTAAGTTAT	TTTTCTCTTATTAATAATAAAAAACATTTAGAAAAAATAGAT	900		
QY	901	TGCCAGACTAGCCCTAGAA	TATTTTTCCCAATTAATTAACAATCACTGTGTATTAATTTG	960		
DB	901	TGCCAGACTAGCCCTAGAA	TATTTTTCCCAATTAATTAACAATCACTGTGTATTAATTTG	960		
QY	961	GCCAGCGCCCATTAATTA	TTTAAACCGAACTGAAATCGAGCGAAACCAATCTGAGCTAT	1020		
DB	961	GCCAGCGCCCATTAATTA	TTTAAACCGAACTGAAATCGAGCGAAACCAATCTGAGCTAT	1020		
QY	1021	TTTCTCTAGATTAATTAAT	TAAGGAGAGAGAGAAAGAAATCAAGTTTAAAGTATGTGTCC	1080		
DB	1021	TTTCTCTAGATTAATTAAT	TAAGGAGAGAGAGAAAGAAATCAAGTTTAAAGTATGTGTCC	1080		
QY	1081	TGAGATGTGCGGTTTGGCA	CGATAGACACCGTATCAATAGCTCATAGGTGCTAGGTGCTAGCTCA	1144		
DB	1081	TGAGATGTGCGGTTTGGCA	CGATAGACACCGTATCAATAGCTCATAGGTGCTAGGTGCTAGCTCA	1144		
QY	1141	GGTTTCGAGAGCTCTGTG	TCACTTCAATGAGCATACTA	CATGCTTGTTCAACCGTTGCTC	1200	

Db 1141 GGTTCGGGACGCTCTCGGTGATCTCACATGCGATCTGCTTGTTCACCGCTTCGTC 1200
Qy 1201 TTGTCCATGCTCCAGACCTTCCTATTTCTGAAACCAAGAGATACCTACTCCCAACAT 1260
Db 1201 TTGTCCATGCTCCAGACCTTCCTATTTCTGAAACCAAGAGATACCTACTCCCAACAT 1260
Qy 1261 CCATCTTACTCATGCACTTCGATGCAACAGCAGCATATGTTTCTCGAAC 1311
Db 1261 CCATCTTACTCATGCAACCTTCGATGCAACAGCAGCATATGTTTCTCGAAC 1311
RESULT 5
AF360356 3343 bp DNA linear PLN 12-MAY-2001
LOCUS AF360356
DEFINITION Zea mays male fertility protein (M45) gene, complete cds.
VERSION AF360356.1 GI:14028756
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 3343)
Fox, T.W., Trimmell, M.R. and Albertsen, M.C.
Cloning of M45, a gene required for male fertility from Zea mays
Unpublished
REFERENCE 2 (bases 1 to 3343)
Fox, T.W., Trimmell, M.R. and Albertsen, M.C.
Submitted (13-MAR-2001) Trait and Technology Development, Pioneer
Direct Submission
Submitted (13-MAR-2001) Trait and Technology Development, Pioneer
Hi-Bred Inc., Inc., 7300 N.W. 62nd Ave., P.O. Box 1004, Johnston,
IA 50131-1004, USA
FEATURES
source
Location/Qualifiers
1. .3343
/organism="Zea mays"
/mol_type="genomic DNA"
/culivar="B73"
/db_xref="taxon:4577"
/chromosome="9L"
<1392..>2942
/gene="M45"
join(1392..1768,1898..2182,2280..2447,2534..>2942)
/gene="M45"
/product="male fertility protein"
join(1392..1768,1898..2182,2280..2447,2534..>2942)
/gene="M45"
/codon_start=1
/product="male fertility protein"
/protein_id="AKS2489.1"
/db_xref="GI:14028757"
/translation="MEKRNLOMRGRDGIQVPHLFEALALALVADPGSLAEV
DVRPVKBLAPGEVWGSMRDNASRLRGLEFGEVGEVPSIEEDLOGRPVAGLA
DEGVAVKMGSEKMETPAVKNPMSEEVCAVGNSTTRKQHEKEECGRPLGRFGE
TGLVYADAVYGLVAVQSGVAVSVAARELDGPIRIPANDLVHNGSVFPTDSMKY
SRDHLNILLBEGGTGLRLLYDEPSTGVHVLKGLVFPNGVOISDHOGLFSEITNC
RIRWYLEGRAGEVEVFANLPGFNDNVNSNGRGQFVAIDCERTPAQGVFAFRPLR
TLVFKPPLSLKVLTKAKARMHMTVLALLDDEGRVVEVLDEDRGHEVVKLVSEVREVRK
LWIGTVAHHNIATIPYLED"

ORIGIN
Query Match 100.0%; Score 1311; DB 8; Length 3343;
Best Local Similarity 100.0%; Pred. No. 3.5e-252;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CCATGCTGCTCTATGAAAAAGATGATGATGCTATACCGCTTTCTTAGGCTGC 60
Db 1 CCATGCTGCTCTATGAAAAAGATGATGATGCTATATCCGTTTCTTAGGCTGC 60
Qy 61 CTTCTTCTGCTTATTACTGATGATCGGGTTTACAAAAAATTCCACGGGTGCATGAT 120
Db 61 CTTCTTCTGCTTATTACTGATGATCGGGTTTACAAAAAATTCCACGGGTGCATGAT 120

Qy 121 CTCGATGTTCCACTTCTCCCACTCGCGCTTGCAATTTCTTGGAATGTCGGGTTCAT 180
Db 121 CTCGATGTTCCACTTCTCCCACTCGCGCTTGCAATTTCTTGGAATGTCGGGTTCAT 180
Qy 181 CTGACCGAGGCCCATGACACCTTTGCGGACACCCATGAGGACCTTTCGATGCGACCA 240
Db 181 CTGACCGAGGCCCATGACACCTTTGCGGACACCCATGAGGACCTTTCGATGCGACCA 240
Qy 241 CGAGAGTATCGGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
Db 241 CGAGAGTATCGGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
Qy 301 TTATATTTCTTATGATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 360
Db 301 TTATATTTCTTATGATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 360
Qy 361 GGGCTCAGCATGATGATTTTGGCTTGGGCGGAGAAATGCGAGGACCGCATGCTAGTG 420
Db 361 GGGCTCAGCATGATGATTTTGGCTTGGGCGGAGAAATGCGAGGACCGCATGCTAGTG 420
Qy 421 TCCACTATTTGGCACTACCGAGAACAGATTTAAAAAAATTAACCAAGTAACTAATCACT 480
Db 421 TCCACTATTTGGCACTACCGAGAACAGATTTAAAAAAATTAACCAAGTAACTAATCACT 480
Qy 481 CGAAGCTATCATGTAATGTTTAAAGAAACATCTATTTAAACCGATCTCTTAAAAA 540
Db 481 CGAAGCTATCATGTAATGTTTAAAGAAACATCTATTTAAACCGATCTCTTAAAAA 540
Qy 541 CAAGCATATTTGGAAGAGACAAATTAAGTTACGTTTCAAAACATTAAGCGGACAA 600
Db 541 CAAGCATATTTGGAAGAGACAAATTAAGTTACGTTTCAAAACATTAAGCGGACAA 600
Qy 601 TTATATCGAAGGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 660
Db 601 TTATATCGAAGGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 660
Qy 661 ATTTGTTTTATATACATTTTCTCTTTCAATAGATATTTCTCGATTTTAA 720
Db 661 ATTTGTTTTATATACATTTTCTCTTTCAATAGATATTTCTCGATTTTAA 720
Qy 721 ATGACTATTAAGCATTTTATTAAGACAGCAGATGCTGATTCGTTCAAAAATC 780
Db 721 ATGACTATTAAGCATTTTATTAAGACAGCAGATGCTGATTCGTTCAAAAATC 780
Qy 781 TTTCTGATTTTAAAGCTAGTTTGGCAACCTGTTCTTCAAGAAATTTGATTT 840
Db 781 TTTCTGATTTTAAAGCTAGTTTGGCAACCTGTTCTTCAAGAAATTTGATTT 840
Qy 841 TTCAAAAAAATTAAGTTATTTTCTTTTATTAATAATGAACACTTAGAAAAATGAGT 900
Db 841 TTCAAAAAAATTAAGTTATTTTCTTTTATTAATAATGAACACTTAGAAAAATGAGT 900
Qy 901 TCCGAGATAGCCCTAGAAATGTTTCCCAATTAATTAAGATCACTGATTAATTAATTG 960
Db 901 TCCGAGATAGCCCTAGAAATGTTTCCCAATTAATTAAGATCACTGATTAATTAATTG 960
Qy 961 GCGAGCCCATTAATTAATTAACGAACTGAATCGAGGAAACCAATCTGAGCTAT 1020
Db 961 GCGAGCCCATTAATTAATTAACGAACTGAATCGAGGAAACCAATCTGAGCTAT 1020
Qy 1021 TTCTCTAGATTAAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Db 1021 TTCTCTAGATTAAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Qy 1081 TGAGATGTCGGGTTTGCAACATGACCGATTAATCAAGTCAATAGGTGCTAGGTCA 1140
Db 1081 TGAGATGTCGGGTTTGCAACATGACCGATTAATCAAGTCAATAGGTGCTAGGTCA 1140
Qy 1141 GGTTCGGGACGCTCTGCTGATCTGATGATGATGATGATGATGATGATGATGATGAT 1200
Db 1141 GGTTCGGGACGCTCTGCTGATCTGATGATGATGATGATGATGATGATGATGATGAT 1200

Measuring, J., Abouelella, A., Allen, N., Anderson, M., Anderson, S., Arachi, H. M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boutgangalter, B., Camarata, J., Chang, J., Chepel, Y., Collymore, A., Cook, A., Cooke, P., Cornu, B., Dextralano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Fato, S., Ferrelle, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Haggopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kemat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbutt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J.,

Meunus, L., Mhova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C.,
 Nguyen, T., Nicol, R., Norbu, C., O'Connor, O., O'Donnell, P.,
 O'Neill, D., Oliver, J., Peterson, K., Phunthang, P., Pierre, N.,
 Rachupa, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
 Roman, J., Schauer, S., Schuppach, R., Seaman, S., Severy, P., Smith, C.,
 Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
 Talama, J., Testaye, S., Theodore, J., Topham, K., Travers, M.,
 Vassiliev, H., Venkataraman, V.S., Viet, R., Vo, A., Wilson, B., Wu, X.,
 Wynn, D., Young, G., Zahoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (17-AUG-2004) Whitehead Institute/MIT Center for Genome
 Research 320 Charles Street, Cambridge, MA 02141, USA
 On Aug 17, 2004 this sequence version replaced gi|19658655.

/product="hypothetical protein"
 /protein_id="AAS38625.1"
 /db_xref="GI:42733681"
 /translation="MRLQTLLEILLFOLKSTPSSNDISLTITLSTIPENKDYNNQ
 LLSKEMFELSEYSSNKRYIFEFVDFGNHEYLENDRITKEFRENKYGVDFSPFLC
 NKFNYLLEKKEFTISNNKKYIPENILNLTSETHENYNNSTLNNKNNVEYEVYV
 PEKQYIKSLQNGYNIGIKSSKNTIKFKFENKDKLSLFFPLIKNAENQEDYVYV
 SSPLSWDTFKQFRLKNDLELTISLDEVAVNLCSNNNNNNNVDEDTLLLSGSGG
 SGGSGGSGTSGAQLSGTYVYNHQ"
 join(12957..3173,3272..5583)
 /note="Geneid exon scores (in order of location ranges) :
 24.88, 254.96 - GSCJ_ID dd_00753"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="AAS38626.1"
 /db_xref="GI:42733682"
 /translation="MEFPFKYVYITSNDKPEMTQSEMSLTITPTDIOKSPVNNSG
 SSSSGKFEATLLDITAGOEYSNARDYVQDGMVPSIDRSSEPEYQSLK
 QHIEVYKDRDVPITIVGNKDLBESRQVSRLEADQLARSIVPIETESKATNITEE
 APTFLVRHPRTNRYVYVVMGGVGGKALIIQFQNFVEYVPTIEDSYRKYVTS
 GLPPLGGSLNKKSSSSSSSSSSSGSGTCLFNKIFEGDKQSPQOAPSPSTIDTGOI
 STNPLEAVLYSYSMNLKSEVPLITGDCVYCGGCVILSRNSNLTGADSPFTWCHEF
 CYSNLSILKEGELPKNDQSYEVYVSPSTSDGSKREBSIIYICIDGSGMGITTE
 EYPSLOSEVMNKKGVKASSGSPYSISLETCVOSIIPMDRLIOYBNRYVLTSES
 DEVMYITQSNSDYGPVYVAGDKLEFDOLIEGSMYDKLPTSGSSDFKAKIKSL
 EYVQSTALGPALVSAALASQKMLSEVICTDGVNGLALIEDLPGLPAQEFKATYV
 KLAQNNSTLITNIGISGHITDGVIGCEBQNTNGITIHLELAREIRKLTQMPMA
 TVEMSI CHAPLEINKYDSKQGLSRVYQCPNNVSLTDLILSSNRNPREFQYV
 FQIQIKYTLDDGVRCLRVASQLOLRTPNTSTNANSTILAMFTQOAKIADQGEY
 MSHRLIKATLTLSICNDPOMEFPNFEVLAEMAPLITGICIKKQKRYEALAD
 DEIQYFKMKNVHKSFBEGRKDLSRKGELEINQYTNKFT"
 complement(5711..6796)
 /note="Geneid exon scores (in order of location ranges) :
 68.61 - GSCJ_ID dd_00752"
 /codon_start=1
 /product="similar to Plasmodium falciparum (isolate 3D7).
 hypothetical protein"
 /protein_id="AAS38627.1"
 /db_xref="GI:42733683"
 /translation="MSSRVKSLIERYNGITIEKSLSPGLINTNINNNNNNNNN
 GITTTTGGNNNNNSTPSTNPVSFDVGVSHNNNOIYVIDSFVNIINNNNNNNN
 NNN
 KENNNNNYFDESTNEMITTLNSCTLEIGITGNDCSINTNPSTSSSTSLBLNN
 KNSNMMWPSLKLINGINDNNNNISNINFPVIOSSSSLSSTSCSYQSEBHEILOE
 QRLKEQLOKQIDDDDDDDDDDDDDSSPPPPPTTTTFCLSRPPIIKKEENID
 QNSMPPPLDGRSRDLPIPTIANI"
 join(8924..9076,9236..10573)
 /note="Geneid exon scores (in order of location ranges) :
 10.37, 128.60 - GSCJ_ID dd_00729"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="AAS38628.1"
 /db_xref="GI:42733684"
 /translation="MSSSTITPLATTTNNNSSTNGCASHSKTRTLPKXIKNIPNS
 ENNAIYIKKNSYIGENGVNNNNKORCLNSDNGENFEKKNLSTPDSKSNVSTTRKS
 SSSSLKANEYDIESIKKELKSMDOI STGIGSPSSPSRESPISTIDLSITATYK
 LIDAVVAASSSSSSKYITDYLSEKSNATRPLOIDIVESHQEIYVPLSLPTTPYK
 NNNNNNNNNNSOHNNTANGOKYPIFSQIOPFKIAVDSHGTENIINNNNNNNITP
 TKKASNTITPQSNRNHNHNSKTSKONPLPSLNTATPQKQKTTTPBKSCTT
 TTKKSKINDYNAPAAVANSSTTTEBPVALIIGATPTSTNESKPTPLKEDPRLA
 AAQRQIDIGKTAGYKNTIALVPKSKRRPTPKTPNKQVCSKRSWQQLKWRQQLH
 ENYDWTGIPENKEVEASNNSTESTSEKEMNKLIEISEENENN"
 complement(10972..11979)
 /note="Geneid exon scores (in order of location ranges) :
 23.29 - GSCJ_ID dd_00728"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="AAS38629.1"
 /db_xref="GI:42733685"
 /translation="MEKIKYFETIIFDQNKQKIOERTSLIGLKEESFTPEPIYV
 ERYLWQTSIESILKNGTINKSIFBEIYSYCYIGLEFLEPEFYVFLKFTTGGIIL
 DYIEKINSLMMKNEKELNSLIYNNNNKEGEEWELIINEQNTLHKSFEYR
 IINANSIWKSSIDRNISNPSINRSFNFYFEKRSADAGDILITSMIGIMDNY

[illegible]

SPSSSSSTSPSCLSSSSSSTPSQSEIAENKODINNNCNTNGNFESEEBEERKXV
EEOSEIEFEIESFEFEKVEVEFEPEFELPISIEBEIKETSLSLSLIV
DMNDDKSTSSQSSSSSSSSSPSPKIKNDNNNTTTTITTTTITTTTSTKATSS

Query Match 5.4%; Score 73.4; DB 3; Length 25769;

Best Local Similarity 49.4%; Pred. No. 0.00015;

Matches 217; Conservative 0; Mismatches 221; Indels 1; Gaps 1;

```
Qy 438 CCGAACAAGATTAAAAAATTAACCAAGTAATCACTGCAAGCTATCATGTAA 497
Db 2392 CGAACAACACACTAAAAAATTTGTCGAGAAAAAATTTTAAAAA 2451
Qy 498 TGTAAAGAACATCATTAATAACACGATCTCTTAAAAAAACAAGATTTTCAAG 557
Db 2452 TATGAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAA 2511
Qy 558 AGACAAATTATGTACAGTTTACAAACATCTAAGCGACAAATTATGAAAGTAA 617
Db 2512 AATTTATTTATTAATTTATTAATTTATTTATTTATTTATTTATTTATTTAT 2571
Qy 618 CTATGACGTTACAGATTTTCTTTTCACTCTGTATTTTGTATTTTATATACAT 677
Db 2572 ATAAAAATTTATTTATTTATTTAA-ATTTATTTCTATTTTATTTTATTTGA 2630
Qy 678 TTTCTCTCTTACATAGATGATTTTCTCCGATTTTAAATGACTATTAAGTCAT 737
Db 2631 TAAATTTTAAAGATATTTGTTTTTTTTTTTTTTTTTTTTTAAAGTATCAAT 2690
Qy 738 TTTATTAAGACGACCATGTCGTAGATTTCTGTCGTAATAATCTTCTGATTTTAA 797
Db 2691 TTTAAATTTGGGTTTGAATGAAAAATTTCTTTAAATTTTTCAGAAAAATCTTT 2750
Qy 798 AGCTAGTTGGCAACCTGTTCTTCAAGAAATTTGATTTTCAAAAAAATTAAGT 857
Db 2751 TTTTTCATTAACACCAACCAATTTTTCATTTTTCATTTTTCATTTTTCAT 2810
Qy 858 TATTTCTCTTATAAAT 876
Db 2811 TTAATTTTTCATTTTAT 2829
```

RESULT 10

AF034389/c

LOCUS AF034389 713 bp DNA linear INV 04-FEB-1999

DEFINITION Plasmodium falciparum sexual stage antigen (816) gene, promoter and

partial cde.

ACCESSION AF034389

VERSION AF034389.1 GI:3098290

KEYWORDS Plasmodium falciparum (malaria parasite P. falciparum)

SOURCE Plasmodium falciparum

ORGANISM Plasmodium falciparum

REFERENCE Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

AUTHORS Decherling, R.J., Kaan, A.M., Mbacham, W., Wirth, D.F., Eling, W.,

Konings, R.N. and Stunnenberg, H.G.

TITLE Isolation and functional characterization of two distinct

sexual-stage-specific promoters of the human malaria parasite

Plasmodium falciparum

Mo1. Cell. Biol. 19 (2), 967-978 (1999)

JOURNAL 99108072

MEDLINE 9891033

REFERENCE 2 (bases 1 to 713)

AUTHORS Decherling, R.J., Kaan, A.M. and Konings, R.N.H.

TITLE Direct Submision

JOURNAL Submitted (12-NOV-1997) Molecular Biology, University of Nijmegen,

Toernooiveld 1, Nijmegen 6525 ED, The Netherlands

FEATURES

source 1..713

gene /organism="Plasmodium falciparum"

/mol_type="genomic DNA"

/strain="NF54"

/db_xref="taxon:5833"

promoter 1..536
/gene="816"
537..>713
/gene="816"
537..710
/gene="816"
711..>713
/gene="816"
/note="P816"

CDS
/product="sexual stage antigen"
/protein_id="AA012581.1"
/db_xref="GI:3098291"
/translation="M"

ORIGIN

Query Match 5.4%; Score 71.2; DB 3; Length 713;

Best Local Similarity 45.8%; Pred. No. 0.00045;

Matches 282; Conservative 0; Mismatches 333; Indels 1; Gaps 1;

```
Qy 444 CAGATTTAAAAAATAACCAAGTAATCACTCGAAAGCTATCATGTATTA 503
Db 681 CAAAAGAAAAAATTTAAAAAATTTAAAAAATTTAAAAAATTTAAATGTAAG 622
Qy 504 AAGAAACATCTATTTAAACACGATCTCTTAAAAAACAAGATTTGAAAGACAA 563
Db 621 AATATTTATCTATGATTAATAAGCTATTTGAAAAATTAATTTATGAAAAATTA 562
Qy 564 ATTAATGTTACATTTACAAACATCTAAGCGACAAATTATTCGAAAGTACATCA 623
Db 561 ATATATGTAATAATTTTTCATTTATTAATAA-ATATATATATATATATATTTT 503
Qy 624 CGTCAGATTTTCTTTTCATTTCTGTATTTGTATTTGTTTATATATATTTCTT 683
Db 502 TTTTTCATTTTATTTATTTATTTATTTATTTATTTATTTTTCATTTTAAAAA 443
Qy 684 CTCTTCAATAGAGGATTTCTCCGATTTTATTAATGACTATTAAGTCAATTTTATA 743
Db 442 TTTTCAGTATTAATAAGTATTTAAACGAAATATGTAATGAAAGAAAGAAATTTG 383
Qy 744 TAAAGCAGCATGTCGTAGATTTCTGTCAAAAATCTTCGATTTTTCGAGCTAG 803
Db 382 TAAATTTTATATTAATAATTAATAAAGAAATATTAATTTTATTAATTAATAAAG 323
Qy 804 TTTGGCAACCTGTTCTTCAAGAAATTTGATTTTTCAAAAAATTTAGTTATTTT 863
Db 322 TTTATTTATTTTATATATATATTTTTCATTTTTCATTTTTCATATATTTTAAAT 263
Qy 864 CTCTTATTAATAATTAAGAAACCTTGAATAATGAGTTGCCAGACTAGCCCTAGA 923
Db 262 ATATATTTTTCATTTTTCATTTTTCATTTTTCATTTTTCATTTTTCATTTTTC 203
Qy 924 TTCCCAATTAATTAATCACTGTCATTAATTTTGGCCAGCCCATTAATTTTAA 983
Db 202 TTTTTCATTTTTCATTTTTCATTTTTCATTTTTCATTTTTCATTTTTCATTTT 143
Qy 984 CCGAATCGAATTCAGCGAAACCAATCTGAGTATTTCTAGATTAGTAAAGGA 1043
Db 142 TAAATTTTATATTTTAAAGTGTATCATTAATGTTTTCATTTTAAAAAATTTAA 83
Qy 1044 GAGAGAGAGAGAA 1059
Db 82 TATCTATATATTAATA 67
```

RESULT 11

DD086962/c

LOCUS DD086962 3576 bp DNA linear INV 23-JUL-2001

DEFINITION Dictyostelium discoideum miga (miga) gene, complete cds.

ACCESSION U86962

VERSION U86962.1 GI:1841871

KEYWORDS

ORIGIN	/number=2
Query Match	5.4%; Score 71.2; DB 3; Length 3576;
Best Local Similarity	47.2%; Pred. No. 0.00043;
Matches 217; Conservative	0; Mismatches 243; Indels 0; Gaps 0;
Oy	412 TGCTAGTGCACATATTGGCACTACCAGAACAGATTAAAAAAATAACCAAAGTAAC 471
Dd	852 TATTTATTTATTTATTTATTTATTTACATTTGTTTTGTATTTTTAAAAAATAAAAAATAAAAA 793
Oy	472 TAATCCACTCGAAGCATATCATGTATGTTTAAAGAAACAATATTAAAACCGACATCT 531
Dd	792 TAAAAAATATAAAAATATAAAAATATAAAAATATAAAAACAACAAATTTAAATTTTAA 733
Oy	532 CTTAAAAAACAGCATTTTCGAAGAAGACAAATTTATGTATACAGTTTACAAACATCTAAG 591
Dd	732 TTTTAAAAAATATAAAAATATAAAAATATAAAAATATAAAAATATAAAAATATAAAAA 673
Oy	592 AGCAGCAATTTATATTCGAAGAGTAAGCTATAGCGTTTCAGATTTTCTTTTCAATCTTG 651
Dd	672 AAAAAAAAAAAAAAAAAATATATATAGGGTGTGAGAAATGTTTTTTTTTTTGTGTG 613
Oy	652 TATTTGTATATGTTTATATATACATTTCTCTCTTACATATAGAGATTTTCTTCCGA 711
Dd	612 TTGTTATTTTTGTTATTTTAAATATTTATTTATATACATATAAAAATATCTTTAAACACAC 553
Oy	712 TTTTATATAAATGACTATTAAGTCAATTTTATATTAAGACGACGATGCTAGATTTCTCGT 771
Dd	552 ACACACACACACACATATGTTTTTTTTTTGTGTTTTTTGTGTTTTTTTATCAACCT 493
Oy	772 TCAAAAATCTTTCTGATTTTATTAAGACTAGTTTGGCAACCTGTTCTTTCAAGAAAT 831
Dd	492 TTCATAGCGCTTTGATTTAGTTGTTTTTTTTTTTTTTTGAATGTTTTTTTGGAAAAAT 433
Oy	832 TTGATTTTTCAAAAAAATTAGTTTATTTTCTCTTAT 871
Dd	432 ATTTTTTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 393
RESULT 12 EX957346/c	
LOCUS	BX957346 149526 bp DNA linear HTG 08-OCT-2004
DEFINITION	Danio rerio clone CH211-117K16, WORKING DRAFT SEQUENCE.
ACCESSION	BX957346
VERSION	BX957346.13 GI:54019944
KEYWORDS	HTG; HTGS PHASE2; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE	Danio rerio (zebrafish)
ORGANISM	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio. 1 (bases 1 to 149526) Beasley,H. Direct Submission Submitted (07-OCT-2004) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: fish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Oct 9, 2004 this sequence version replaced gi:53850295.
REFERENCE	----- Genome Center Center: Wellcome Trust Sanger Institute Center code: SC Web site: http://www.sanger.ac.uk
AUTHORS	Contact: zfish-help@sanger.ac.uk
TITLE	----- Project Information Project name: zc117K16
JOURNAL	----- Summary Statistics Assembly program: XGAP4; version 4.5 Chemistry: Dye-terminator; 100% of reads Consensus quality: 149526 bases at least Q40 Consensus quality: 149526 bases at least Q30 Consensus quality: 149526 bases at least Q20 Insert size: 149526; sum-of-contigs
COMMENT	

Insert size: 149176; 6.5% error; agarose-fp
Quality coverage: 12.41x in Q20 bases; sum-of-contigs Quality
coverage: 12.54x in Q20 bases; agarose-fp

- * NOTE: This is a 'working draft' sequence.
- * This sequence will be replaced
- * by the finished sequence as soon as it is available and
- * the accession number will be preserved.

FEATURES

Source
1. 149526
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="CH211-117K16"
/clone_1fb="CHOKI-211"
1. 149526
/note="assembly_fragment:03115"
clone_end:SP6
vector_side:right
clone_end:T7
vector_side:left"

misc_feature
1. 149526
/note="assembly_fragment:03115"
clone_end:SP6
vector_side:right
clone_end:T7
vector_side:left"

ORIGIN

Query Match 5.4%; Score 70.8; DB 2; Length 149526;
Best Local Similarity 46.8%; Pred. No. 0.00047;
Matches 260; Conservative 0; Mismatches 292; Indels 4; Gaps 1;

QY 488 TATCATGTAATGTTTAAAGAACATCTATTAAACACGATCTCTTAAAAACAGCAT 547
DB 109063 TTTTAAATATGAAACAAAATCTATATTTTACCTTGGCCCTTGATAGCAT 109004
QY 548 ATTTCGAAAGACAATTATGTTAC- - - GTTTACAAATCTTAAGCCGCAATT 603
DB 109003 GTTTTAAAGATTAAATAATTTTAAACATGTTTATTCACATATATGACTAAGAG 108944
QY 604 TATCGAAAGGTAGCTATGACGTTGAGATTTTCTTTTCAATCTGTTATTTGTTAT 663
DB 108943 CCAAAAGTACTAATGACATTTTAAATGTTTATTTTTCGATTTTATTTTATTT 108884
QY 664 GTTTTATATATACATTTCTCTCTTACATATGATGATGATTTTCTTCGATTTTAAAT 723
DB 108883 AT 108824
QY 724 ACTATTAAGTCAATTTTATATTAAGACACCATGCTGATGATTTCTGTTCAAAATCTTT 783
DB 108823 TATATATTTTAAATTTATATATTAATAATTTATGTAATTTAAACATCAATTAC 108764
QY 784 CTGATTTTAAAGTAGTGTGGCAACCTGTTCTTCAAGAATTTTGAATTTTTC 843
DB 108763 ATTATATCATTTTAAACCTCTTATATGTTAATGTTATATATATATATATATAT 108704
QY 844 AAAAAAATAGTTTATTTTCTTTATTAATAATAGAAAACATCTAGAAAAATAGAGTTGC 903
DB 108703 AGAAAAACAAAGTATATATTTATTTATTTATTTTAAATTAATAATTTT 108644
QY 904 CAGACTAGCCCTGGAATGTTTCCCAATTAATTAATCAATCCTGTATATTTTGGCC 963
DB 108643 TATATATGCCCTTCATTAGCATCTTATTTTAAAAATATATATATATATTTTGGTC 108584
QY 964 AGCCCATTAATTTTAAACGAAACTGAATCGAGCAACCAATCTGACCTATTC 1023
DB 108583 ACATATATATATTTAAGAGACCAAAAAGTAAATAATATGACTTATTTTATATCTGAA 108524
QY 1024 TCTAGATTAAGTAAAA 1039
DB 108523 GTTTTATAGTTAAAA 108508

RESULT 13
PFMAL4P1/c PFMAL4P1 347582 bp DNA linear INV 29-JAN-2003
DEFINITION Plasmodium falciparum MAL4P1.
ACCESSION AL034557 AL844503

VERSION

AL034557.8 GI:23498126

KEYWORDS

Plasmodium falciparum 3D7

SOURCE

Plasmodium falciparum 3D7

ORGANISM

Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

REFERENCE

1 Hall, N., Pain, A., Berriman, M., Churcher, C., Harris, B., Harris, D., Mungall, K., Bowman, S., Atkin, R., Baker, S., Barton, A., Brooks, K., Buckee, C.O., Burrows, C., Cherevach, I., Chillingworth, C., Chillingworth, T., Christodoulou, Z., Clark, L., Clark, R., Corton, C., Cronin, A., Davies, R., Davis, P., Dear, P., Dearden, F., Doggett, J., Felwell, T., Goble, A., Goodhead, I., Gilliam, R., Hamlin, N., Hance, Z., Harper, D., Hauser, H., Hornby, T., Holroyd, S., Horrocks, P., Humphray, S., Jagels, K., James, K.D., Johnson, D., Kerhornou, A., Knights, A., Konfortov, B., Kyes, S., LaRke, N., Lawson, D., Lennard, N., Line, A., Maddison, M., McLean, J., Mooney, P., Moulé, S., Murphy, L., Oliver, K., Ormond, D., Price, C., Quail, M.A., Rabinowitch, E., Rajandream, M.A., Ruter, S., Rutherford, K.M., Sanders, M., Simmonds, M., Seeger, K., Sharp, S., Smith, R., Squares, R., Squares, S., Stevens, K., Taylor, K., Tiley, A., Unwin, L., Whitehead, S., Woodward, V., Suleston, J.B., Craig, A., Newbold, C. and Barrell, B.G.

AUTHORS

Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13

TITLE

Nature 419 (6906), 527-531 (2002)

JOURNAL

2 (bases 1 to 347582)

PUBLISHED

22255708

MEDLINE

12368867

REFERENCE

2 (bases 1 to 347582)

AUTHORS

Hamlin, N., Pain, A., Berriman, B., Hall, N., Bowman, S., Churcher, C., Harris, B., Harris, D., Lawson, D., Quail, M. and Barrell, B.

TITLE

Direct Substitution

JOURNAL

Submitted (24-FEB-1999) P.falciparum Genome Sequencing Consortium, The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK

COMMENT

On Oct 3, 2002 this sequence version replaced gi:5731897. For more information about this sequence or the Malaria Project, see http://www.sanger.ac.uk/Projects/P_falciparum.

FEATURES

Location/Qualifiers

1. 347582
/organism="Plasmodium falciparum 3D7"
/mol_type="genomic DNA"
/isolate="3D7"
/db_xref="taxon:36329"
/chromosome="4"
/clone="MAL4P1"
10384. 32958
/note="rep20 Repeats"
join(35153. 41725,42757. 44124)
/gene="VAR"
/note="synonym: PED0005W"
join(35153. 41725,42757. 44124)
/gene="VAR"
/note="Similar to Plasmodium falciparum variant-specific surface protein var-2 SMALL.026033 (EMBL:140609) (2664 aa) fasta scores: E(): 7.8e-149, 43.1% id in 2872 aa"
/codon_start=1
/product="erythrocyte membrane protein 1 (PEEMP1)"
/protein_id="CAD49094.1"
/db_xref="GI:23498127"
/db_xref="GOA:O81220"
/db_xref="UniProt/TREMBL:O81220"

/translation="NVRVSGSGGSGVDVDDIGKNAKHLDSIGKKVHDVQNGADCTGASGDAKNTIDLDKEDLQKAPINIFKLIGTDPCKLVEDYINNVHNGDDKGRFYFCTE LSGKRFQNFPSDTLGGQCTNSKMRSGCBACAPYRRLHCHNLDSLETTSTYADTL LLEVCAAAKYEGQSINTHTYTRKHSHNKSPSQLCVTLARSPADIDIVAGKDLFYGNT YESAREKLENTLKEVFGKIHGGLSEAAKKTQODGNYVQLRBDMMVTPTRVYKAI TCEVSKGNNVPATGDEKNPSLTSKQRCDDKXKPIKSGGNNVIPTVYDVPQY LMEFEYMAEDRCRLKXKTKDAIKCRKNGEKYCDLNRDKXTLASGKYHFFPEDF CKDCQYSCAPFDVMDNQKLEFLKQKRTYKETSISGSGGSGSGSKRKKDDTTNYSIEY KKFYELKSTYKIVVNNPLEKLNDEVCITKNDIDGNDIPEKVVHSSAKKGDNNK TTYRKYCEACPCWGAKEVGGWAKENKSGQTDYDIDPKTTTITLITLGDTRKSDNVQ KYKFCNGNGGNGEASATPNAITSREKKGDMEMKI CYDENKEKTKGSDAINTCYL ODGKHREOKVTSYNAPFWKVDHMLHDSVBERLNSCTINNAKSNCKNEKCKNKE CGCFEKWKQKKKEKEMEAIKOHFKQXDIIEBTGCDAGVTAAVTKLEFLNEDTIEKS


```

Db      335424  TTTTCAGTATTAAGATTTTAACGAAATATGAGAAATGAAAGAAATTTGGTATG 335365
Qy      744    TAAAGACGCGATGTCGTAGATTCCTGTCGAAAAATCTTTCGATTTTTTTAAGAGCTAG 803
Db      335364  TAAATTTTAAATTAATAAATAAAGAAATATATATTTTATTAATTAATTAATAAAG 335305
Qy      804    TTGGCAACCCGTTCTTTCACAAAGAAATTTGATTTTTTTCAAAAAAATTAAGTTATTTT 863
Db      335304  TTTATTTATTTTATATATATTAATTTTAAATTTTAAATTCATATATTTTAAATTT 335245
Qy      864    CTCCTTATTAATAAGAAAAACCTTAGAAAAATAGAGTGCAGACTGACCCCTAGAAATGT 923
Db      335244  ATATTTTTTTTTTTTTTTTTCCTTATATATATATATATATAGCTAATATATATATTT 335185
Qy      924    TTCCTAATAATTAACAATCACTGTATTAATTTTGGCAGCCCAATAATTAATTTAA 983
Db      335184  TTTTTTTTTTTTTTTTAAATTAATATATTTATTTTTCACATATATATAATTAAT 335125
Qy      984    CCGAACTGAAATCGAGCGAAACCAATCTGAGCTATTTCTTAGATTAGTAAAGGCA 1043
Db      335124  TAAATTTTATTTTAAAGGTGATCATTAATGTTTTTAAATTAATAAAAAAATAGT 335065
Qy      1044  GAGAGAGAGGAGAAA 1059
Db      335064  TATCTATATTAATAAA 335049

```

```

RESULT 14
LOCUS      CQ422641      883 bp      DNA      linear      PAT 28-JAN-2004
DEFINITION Sequence 7675 from Patent WO0151628.
ACCESSION  CQ422641
VERSION    CQ422641.1  GI:41374870
KEYWORDS
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE  1. 883
AUTHORS    Lillie, J., Xu, Y., Wang, Y. and Steinmann, K.
TITLE      Novel gene, compositions, kits, and methods for identification,
JOURNAL    Patent: WO 0151628-A 7675 19-JUL-2001;
           Millennium Pharmaceuticals, Inc. (US)
FEATURES
   source   1..883
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
ORIGIN
Query Match      5.4%; Score 70.2; DB 6; Length 883;
Best Local Similarity 39.2%; Pred. No. 0.0007;
Matches 237; Conservative 0; Mismatches 367; Indels 0; Gaps 0;

```

```

Qy      419  TGTCCACTATTGGCACTACCCAGACAAAGATTTAAAAAATACCAAGTAAGTAATCCA 478
Db      873  TGTNNNNNNNANACCTTAAACCTTAAANATTTNNNANNTNANNAAAAAAATATTTAANN 814
Qy      479  CTCGAAGCTATGTAATGTTTAAAGAAACATCTATTAAAAACAGATCCCTTTAAA 538
Db      813  TAAAAATTTTAAAAAATCTAATTTNANNTANNTAATTTTNTAAAAAANNAAAAA 754
Qy      539  AACAGCATATTTGAAAGAGACAAATTTATGTTACAAACATCTTAGAGGACGA 598
Db      753  TTTAAANNNTTNTNANTATTAACCAAAATTTTTTTTTTAAAAAAATTTTTTAAANNT 654
Qy      599  AATTATATCGAAAGTAAGCTATGAGCTGATGATTTTCTTTTTCATTTCTGTTATTTG 658
Db      693  AATTAATTTNAATTTTTTTTATATNAAAAAAATTTTNAATTTTANCAANTNTTTT 634
Qy      659  TTATTTGTTTATATACATTTTCTCTTACATAGAGATTTTCTTCGATTTTATA 718

```

```

Db      633  TTTNTTNNATTAATAAAAAATTTTAAATTAATAAANANNTTTTTTTATTAANAATATTAA 574
Qy      719  AAATGCTATTAAGCTATTTTATATATAGACGCGATGTCGTAGATTCGTTCAAAA 778
Db      573  AAAAAAANNNTTTTTTTTTTATANTTTATTAATAAATAATTTATNTNTTTTTTCNAAAA 514
Qy      779  TCTTTGATTTTTTTTAAAGCTAGTTGGCAACCTGTTCTTTCAAGAAATTTGATT 838
Db      513  AAAAAAATTTTNNAAATNTTTTAAAACTTNAATTAANANNAATTTTNTNNNTT 454
Qy      839  TTTTCAAAAAAATAGTATTTCTTTTATTAATTAATAAGAAACCTTAGAAAAATAGA 898
Db      453  NNNATTTAAAAATTTATTTTNTTTTNTTCTATTAATTAANANANANNTAATAAT 394
Qy      899  GTTCCGACGCTAGCCCTAGAAATGTTTCCCAATTAATTAACAATCTGTATTAATTT 958
Db      393  TTTNAAAAAATATATTTATANTTANNTANNAANANAAATTTAAANTNANANAAAAATTT 334
Qy      959  TGGCCAGCCCATTAATTTTAAACCGAACTGAAATGAGCGAAACCAATCTGAGCT 1018
Db      333  NTTAAATTAATTAATAATTAATTAACNTCAATTTTAAATTTATATTAATAAAAAAAT 274
Qy      1019  ATTT 1022
Db      273  NTT 270

```

```

RESULT 15
LOCUS      BV119878/c      241 bp      DNA      linear      STS 19-MAR-2004
DEFINITION PZA01377 Zea mays CML247 Zea mays STS genomic, sequence
ACCESSION  BV119878
VERSION    BV119878.1  GI:45589251
KEYWORDS   STS.
SOURCE     Zea mays
ORGANISM   Zea mays
REFERENCE  1. 241
AUTHORS    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
TITLE      Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
JOURNAL    1 (bases 1 to 241)
           McMillen, M.D., Vrch Bi, I., Schroeder, S.S. and Gaut, B.S.
           MP2-UCI Joint SNP Discovery
           Unpublished (2003)
COMMENT
Contact: Brandon S. Gaut
Dept. Ecology and Evolutionary Biology
U.C. Irvine
321 Steinhilber Hall, Irvine, CA 92697-2525, USA
Tel: (949) 824-2564
Fax: (949) 824-2181
Email: bgaut@uci.edu
Primer A: TTTCACTGCTCAATGATCAAGA
Primer B: AGAAACCAAGCTTAGAGAGA
STS size: 241
Protocol:
PCR amplification of genomic DNA
Template: 50 ng
Primer: each 0.5 uM
dNTPs: each 200 uM
Tag Polymerase: RedTaq (Sigma)
Total Vol: 10 uL
Amplicon sequencing
ABI protocol - using d-Rhodamine terminator cycle
sequencing ready reaction with amplified DNA polymerase FS
Sequence ran on ABI 3700 sequencer.

```

```

Buffer:
Genomic DNA amplification
RedTaq (Sigma)
Sequencing buffer
d-Rhodamine kit (ABI)

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 08:10:17 ; Search time 634.708 Seconds
(without alignments)
12227.321 Million cell updates/sec

Title: US-10-713-381-2_COPY_1_1311

Perfect score: 1311

Sequence: 1 catggtgctctctatcgaaaa.....cgcacatcgtcttcctgaac 1311

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq.16Dec04:*
1: genebegn19808:*
2: genebegn19908:*
3: genebegn20008:*
4: genebegn20018:*
5: genebegn20018:*
6: genebegn20028:*
7: genebegn20028:*
8: genebegn20038:*
9: genebegn20038:*
10: genebegn20038:*
11: genebegn20038:*
12: genebegn20048:*
13: genebegn20048:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1311	100.0	1394	2	AAx07408	Aax07408 Zea mays
2	1311	100.0	1394	2	AAx07409	Aax07409 Zea mays
3	1311	100.0	1394	5	AAH76332	Aah76332 Z. mays M
4	1311	100.0	1394	5	AAH76333	Aah76333 Z. mays M
5	160.6	12.3	255	5	AAH76340	Aah76340 Z. mays M
6	146	11.1	158	5	AAH76334	Aah76334 Z. mays M
7	70.2	5.4	883	4	AA15210	Aa15210 Human bre
8	70.2	5.4	960	11	ACN85231	Acn85231 Breast ca
9	69.8	5.3	13400	13	ADR04296	Adr04296 Corn FT h
10	67.2	5.1	2657	2	AAZ10551	Aaz10551 DNA beque
11	66.4	5.1	6027	2	AAx58751	Aax58751 Maize dul
12	66.4	5.1	6027	8	ABX09935	Abx09935 DNA encod
13	66.4	5.1	6027	12	ADK12106	Adk12106 cDNA encd
14	66.2	5.0	439	8	ABX35844	Abx35844 Bovine ES
15	65.6	5.0	346	4	AA187279	Aa187279 Human pol
16	65.6	5.0	8056	8	ABZ10246	Abz10246 Haematopo
17	64.6	4.9	9859	13	ADs89714	Ads89714 Oligonuc1
18	64.4	4.9	9859	13	ADs89440	Ads89440 Oligonuc1
19	64	4.9	8056	8	ABZ10100	Abz10100 Haematopo
20	63.8	4.9	431	8	ABx44556	Abx44556 Bovine ES

C	21	63.6	4.9	335913	5	AA161371	Aa161371 Soybean 2
C	22	63.6	4.9	335913	5	AA161372	Aa161372 Soybean 2
C	23	63.4	4.8	12968	4	AA545494	Aa545494 Chemicall
C	24	63.4	4.8	12968	4	AA546779	Aa546779 Tumour su
C	25	63.4	4.8	12968	6	AB134085	Ab134085 Human imm
C	26	63.4	4.8	12968	6	ABK28424	Abk28424 DNA trans
C	27	63.4	4.8	12968	6	ABN80285	Abn80285 Human che
C	28	63	4.8	7624	6	AB134113	Ab134113 Human imm
C	29	62.4	4.8	10762	6	AB192322	Ab192322 Chemicall
C	30	61.4	4.7	9652	6	AB132909	Ab132909 Human imm
C	31	60.6	4.6	565	13	ACN56344	Acn56344 Cotton an
C	32	60.6	4.6	5822	5	AB133097	Ab133097 Human imm
C	33	60.4	4.6	500	5	AD173089	Ad173089 Human ova
C	34	60.4	4.6	500	5	AD138222	Ad138222 Human ova
C	35	60.2	4.6	5818	4	AA546636	Aa546636 Tumour su
C	36	60.2	4.6	8896	6	ABN80326	Abn80326 Human che
C	37	59.4	4.5	461	5	AD143834	Ad143834 Human ova
C	38	59.4	4.5	14919	4	AA546506	Aa546506 Tumour su
C	39	59.2	4.5	19459	6	ABK31213	Abk31213 Signal tr
C	40	59.2	4.5	19459	6	AB170528	Ab170528 Chemicall
C	41	59	4.5	517	13	ACN56273	Acn56273 Cotton an
C	42	59	4.5	5303	6	AB132871	Ab132871 Human imm
C	43	58.8	4.5	393	8	ABX39417	Abx39417 Bovine ES
C	44	58.6	4.5	392	5	AD143653	Ad143653 Human ova
C	45	58.6	4.5	529	13	ACN56029	Acn56029 Cotton an

ALIGNMENTS

RESULT 1
AAx07408
ID AAx07408 standard; DNA; 1394 BP.
XX
XX
AC AAx07408;
XX
DT 08-JUN-1999 (first entry)
XX
DE Zea mays M45 male tissue-preferred regulatory region.
XX
XX M45; male; tissue-preferred; regulatory region; plant cells;
KM plant tissue; differentiated; maize; hybrid seed; fertility; ss.
XX
OS Zea mays.
XX
XX
PN W09859061-A1.
XX
PD 30-DEC-1998.
XX
PF 19-JUN-1998; 98WO-US012895.
XX
PR 23-JUN-1997; 97US-00880499.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman GA, Kendall TL;
XX
DR WPI; 1999-105628/09.
XX
PT New nucleic acid encoding a M45 male tissue-preferred regulatory region
XX - useful in mediating plant fertility, especially hybrid seed production.
XX
PS Claim 2; Page 22-23; 39pp; English.
XX
XX
The sequence is that encoding an M45 male tissue-preferred regulatory
region. It may be used in the construction of a vector for a method of
producing exogenous genes in a male tissue- preferred manner, which is
useful in restoring or conferring fertility, such as in hybrid seed
production. In conferring fertility, a monocot/dicot plant is transformed
with the exogenous nucleotide sequence (a male sterility gene, preferably
Me45), which encodes a product selected from auxins, rolgs and diptheria
toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
and infertile plants

XX Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;
 SQ Query Match 100.0%; Score 1311; DB 2; Length 1394;
 Best Local Similarity 100.0%; Pred. No. 1e-269;
 Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCATGGTGTCTATGAAAAAGATGATACATGTGTCTATATCCGTTTCTTAGGGTCC 60
 DB 1 CCATGGTGTCTATGAAAAAGATGATACATGTGTCTATATCCGTTTCTTAGGGTCC 60

QY 61 CTTCTTCTGCTTATTAAGTGAATCGGGGTTTACAAAACTTCCACGGGTGCATGAT 120
 DB 61 CTTCTTCTGCTTATTAAGTGAATCGGGGTTTACAAAACTTCCACGGGTGCATGAT 120

QY 121 CTCGATGTTCCACTTCCCACTCCGGGTGCACTTTCTGGAATGTGGTGGTCCCAT 180
 DB 121 CTCGATGTTCCACTTCCCACTCCGGGTGCACTTTCTGGAATGTGGTGGTCCCAT 180

QY 181 CTGACCGAGGCCCATCAGACACTTTCGGGACACCCATCAAGGGCTTTTGGATGGCCCA 240
 DB 181 CTGACCGAGGCCCATCAGACACTTTCGGGACACCCATCAAGGGCTTTTGGATGGCCCA 240

QY 241 CGAAGCGTATCGGGTGTGTGATCAAGGGATATATGTCCTCCACATCGTCACTTATA 300
 DB 241 CGAAGCGTATCGGGTGTGTGATCAAGGGATATATGTCCTCCACATCGTCACTTATA 300

QY 301 TTATTAATCTTTAGATATTTATTTTGGAAAAATACAACTTATATCTTTGTGTA 360
 DB 301 TTATTAATCTTTAGATATTTATTTTGGAAAAATACAACTTATATCTTTGTGTA 360

QY 361 GGGCTCAGCATAGATTTTGTAGGGCCAGAAAAATGCGAGGACGACCATCTTAGTG 420
 DB 361 GGGCTCAGCATAGATTTTGTAGGGCCAGAAAAATGCGAGGACGACCATCTTAGTG 420

QY 421 TCCACTATTGGCACTACCCAGAACAAAGATTAAAAAATACCAAGTAACTAATCCACT 480
 DB 421 TCCACTATTGGCACTACCCAGAACAAAGATTAAAAAATACCAAGTAACTAATCCACT 480

QY 481 CGAAGCGTATCATGATATGTTTAAAGAAACATCTATTTAAACACGATCCTCTTAAAAA 540
 DB 481 CGAAGCGTATCATGATATGTTTAAAGAAACATCTATTTAAACACGATCCTCTTAAAAA 540

QY 541 CAAGCATATTTGCAAGAGACAATATATGTTACAGTTTACAAACATCTAAGGCGACAAA 600
 DB 541 CAAGCATATTTGCAAGAGACAATATATGTTACAGTTTACAAACATCTAAGGCGACAAA 600

QY 601 TTATATCGAAAGGTAAAGCTATGACGTTTCAGATTTTCTTTCAATCTTGTATTTGTT 660
 DB 601 TTATATCGAAAGGTAAAGCTATGACGTTTCAGATTTTCTTTCAATCTTGTATTTGTT 660

QY 661 ATTGTTTTTATATACATTTTCTCTTCAATAGATGATTTTCTTCGATTTTATATA 720
 DB 661 ATTGTTTTTATATACATTTTCTCTTCAATAGATGATTTTCTTCGATTTTATATA 720

QY 721 ATGACTATAAAGTCATTTTATATAAGAGCAGCATGTCTAGATTTCTGTTCAAAAAATC 780
 DB 721 ATGACTATAAAGTCATTTTATATAAGAGCAGCATGTCTAGATTTCTGTTCAAAAAATC 780

QY 781 TTTTCTGATTTTTTAAAGAGCTAGTTGGCAACCTGTTTCTTCAAGAAATTTGATTTT 840
 DB 781 TTTTCTGATTTTTTAAAGAGCTAGTTGGCAACCTGTTTCTTCAAGAAATTTGATTTT 840

QY 841 TTCAAAAAAATATAGTTTATTTCTGTTTATAAATAGAAAAACATTAAGAAAAATGAGCT 900
 DB 841 TTCAAAAAAATATAGTTTATTTCTGTTTATAAATAGAAAAACATTAAGAAAAATGAGCT 900

QY 901 TGCCAGACTAGCCCTAGAAATGTTTCCCAATTAATACATCACTGTATATATTTTG 960
 DB 901 TGCCAGACTAGCCCTAGAAATGTTTCCCAATTAATACATCACTGTATATATTTTG 960

QY 961 GCCAGGCCCATTAATTTATTTAAACGAAACTGAATCGAGCAACCAATCTGAGCTAT 1020
 DB 961 GCCAGGCCCATTAATTTATTTAAACGAAACTGAATCGAGCAACCAATCTGAGCTAT 1020

DB 961 GCCAGGCCCATTAATTTATTTAAACGAAACTGAATCGAGCAACCAATCTGAGCTAT 1020
 QY 1021 TTCTCTAGATTAGTAAAGAGAGAGAGAGAAAGAAATCACTTTTAACTATGTCCTCC 1080
 DB 1021 TTCTCTAGATTAGTAAAGAGAGAGAGAGAGAAAGAAATCACTTTTAACTATGTCCTCC 1080

QY 1081 TGAAGATGTGCGGTTTGGCAACATAGCCACCGTAAATCATAGCTCATAGGTGCTACGTA 1140
 DB 1081 TGAAGATGTGCGGTTTGGCAACATAGCCACCGTAAATCATAGCTCATAGGTGCTACGTA 1140

QY 1141 GGTTCGGACGCTCTCGTGTATCTCATGATGATGATCATATGCTTGTTCACCGTTGTC 1200
 DB 1141 GGTTCGGACGCTCTCGTGTATCTCATGATGATGATCATATGCTTGTTCACCGTTGTC 1200

QY 1201 TTGTTCCATGCTGCAAGCCCTTTCCTTATGTAACCAAGAGATACCTACTCCCAACAT 1260
 DB 1201 TTGTTCCATGCTGCAAGCCCTTTCCTTATGTAACCAAGAGATACCTACTCCCAACAT 1260

QY 1261 CCATCTTACTCATGCAACTTCCATGCAACGACATATGTTTCTGTAAC 1311
 DB 1261 CCATCTTACTCATGCAACTTCCATGCAACGACATATGTTTCTGTAAC 1311

RESULT 2

AAK07409

ID AAK07409 standard; DNA; 1394 BP.

XX AAK07409;

XX 08-JUN-1999 (first entry)

XX Zea mays Me45 male tissue-preferred regulatory region.

XX Ms45; male; tissue-preferred; regulatory region; plant cells;

XX Plant tissue; differentiated; hybrid seed; fertility; ss.

XX Zea mays.

XX MO9859061-A1.

XX 30-DEC-1998.

XX 19-JUN-1998; 98WO-US012895.

XX 23-JUN-1997; 97US-00880499.

XX (PION-) PIONEER HI-BRED INT INC.

XX Albreten MC, Fox TW, Garmaat CW, Huffman GA, Kendall TL;

XX WPI; 1999-105628/09.

XX New nucleic acid encoding a Me45 male tissue-preferred regulatory region.
 PT - useful in mediating plant fertility, especially hybrid seed production.

XX Claim 3; Page 23-24; 39pp; English.

XX The sequence is that encoding an Ms45 male tissue-preferred regulatory
 CC region. It may be used in the construction of a vector for a method of
 CC producing exogenous genes in a male tissue-preferred manner, which is
 CC useful in restoring or conferring fertility, such as in hybrid seed
 CC production. In conferring fertility, a monocot/dicot plant is transformed
 CC with the exogenous nucleotide sequence (a male sterility gene, preferably
 CC Ms45), which encodes a product selected from auxins, rolb and dipteria
 CC toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
 CC and infertile plants

XX Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;

XX Query Match 100.0%; Score 1311; DB 2; Length 1394;
 XX Best Local Similarity 100.0%; Pred. No. 1e-269;
 XX Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 181 CTGACCGAGGCCCATCAGACACTTTGGGGACACCATCAAGGGCTTTGGGATGGCCA 240
 DB 181 CTGACCGAGGCCCATCAGACACTTTGGGGACACCATCAAGGGCTTTGGGATGGCCA 240
 QY 241 CGAGACGTATCGGCTGTGTGATCGAGGGGATATATGTCCCAATCGTCACTATA 300
 DB 241 CGAGACGTATCGGCTGTGTGATCGAGGGGATATATGTCCCAATCGTCACTATA 300
 QY 301 TTATATCTTTAGATATTTATTTAATTTTGGAAAAATTAACAATTTATCTTTGTGTA 360
 DB 301 TTATATCTTTAGATATTTATTTAATTTTGGAAAAATTAACAATTTATCTTTGTGTA 360
 QY 361 GGGCCCTCAGCATGATTTTGGCTTAGGGGCCAGAAAAATCGAGGACGACATGCTAGTG 420
 DB 361 GGGCCCTCAGCATGATTTTGGCTTAGGGGCCAGAAAAATCGAGGACGACATGCTAGTG 420
 QY 421 TCCACTATTTGGCACTACCCAGAACAAAGATTTAAAAAATTAACAAAGTTAATCTAATCACT 480
 DB 421 TCCACTATTTGGCACTACCCAGAACAAAGATTTAAAAAATTAACAAAGTTAATCTAATCACT 480
 QY 481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTTAAACCAAGATCCTTTAAAAAA 540
 DB 481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTTAAACCAAGATCCTTTAAAAAA 540
 QY 541 CAAGCATATTTCCAAAGAGACAAATTTATGTTACAGTTTACAAACATCTAAGAGGACAAA 600
 DB 541 CAAGCATATTTCCAAAGAGACAAATTTATGTTACAGTTTACAAACATCTAAGAGGACAAA 600
 QY 601 TTATATCGAAAGGTAAAGCTATGACGTTACAGATTTTCTTTTCAATCTTTGTAATTTGTT 660
 DB 601 TTATATCGAAAGGTAAAGCTATGACGTTACAGATTTTCTTTTCAATCTTTGTAATTTGTT 660
 QY 661 ATTGTTTTTATATACATTTTCTTCTCTTCAATAGAGATTTTCTTCCGATTTTATATA 720
 DB 661 ATTGTTTTTATATACATTTTCTTCTCTTCAATAGAGATTTTCTTCCGATTTTATATA 720
 QY 721 ATGACTATTAAGTCAATTTTATATATAGAGACGACGATGCTAGATTCTGCTTAAAAATC 780
 DB 721 ATGACTATTAAGTCAATTTTATATATAGAGACGACGATGCTAGATTCTGCTTAAAAATC 780
 QY 781 TTTCTGATTTTAAAGAGCTAGTTTGGCAACCTGTTCTTTCAAGAATTTTGAATTT 840
 DB 781 TTTCTGATTTTAAAGAGCTAGTTTGGCAACCTGTTCTTTCAAGAATTTTGAATTT 840
 QY 841 TTTCAAAAAAATTAAGTTATTTCTCTTTAATAAATGAAAAACAATTAGAAAAATAGAGT 900
 DB 841 TTTCAAAAAAATTAAGTTATTTCTCTTTAATAAATGAAAAACAATTAGAAAAATAGAGT 900
 QY 901 TGGCAGACTAGCCCTGAATGTTTCCCAATAAATTAATCACTGCTGTAATTAATTTTG 960
 DB 901 TGGCAGACTAGCCCTGAATGTTTCCCAATAAATTAATCACTGCTGTAATTAATTTTG 960
 QY 961 GCCAGGCCCATTAATTTTAAACCGAACTGAATGAGCGAAACCAATCTGAGCTAT 1020
 DB 961 GCCAGGCCCATTAATTTTAAACCGAACTGAATGAGCGAAACCAATCTGAGCTAT 1020
 QY 1021 TTCTCTAGATTTAGTAAAAAGGAGAGAGAGAGAAATCAAGTTTAATCAATGTC 1080
 DB 1021 TTCTCTAGATTTAGTAAAAAGGAGAGAGAGAGAAATCAAGTTTAATCAATGTC 1080
 QY 1081 TGAAGATGTGGGTTTGGCAACGATAGCACCCTGTAATCAAGCTCAGAGCTCAGTCA 1140
 DB 1081 TGAAGATGTGGGTTTGGCAACGATAGCACCCTGTAATCAAGCTCAGAGCTCAGTCA 1140
 QY 1141 GGTTCGAGACTCTGCTGCTATCTCAATGAGCAATGCTGTTTCAACGTTGCT 1200
 DB 1141 GGTTCGAGACTCTGCTGCTATCTCAATGAGCAATGCTGTTTCAACGTTGCT 1200
 QY 1201 TTGTTTCATGCTCGAAGCTTTGCTTATCTGAAACCAAGAGAGATCTCTCCAAACAT 1260
 DB 1201 TTGTTTCATGCTCGAAGCTTTGCTTATCTGAAACCAAGAGAGATCTCTCCAAACAT 1260
 QY 1261 CCATCTTACTACGCAACTTCATGCAACGCGACATATGTTCCGAAC 1311

DB 1261 CCATCTTACTACGCAACTTCATGCAACGCGACATATGTTCCGAAC 1311

RESULT 5
 AAH76340
 ID AAH76340 standard; DNA; 255 BP.
 XX
 AC AAH76340;
 XX
 DT 29-OCT-2001 (first entry)
 XX
 DE Z. mays Ms45 promoter fragment.
 XX
 XX Ms45; male tissue; regulatory region; transcription; male fertility;
 KW hybrid seed; promoter; ds.
 OS
 OS Zea mays.
 PN W0200160997-A2.
 XX
 XX 23-AUG-2001.
 PD
 XX 13-FEB-2001; 2001WO-US04527.
 PF
 XX 15-FEB-2000; 2000US-00504487.
 PR
 XX (PION-) PIONEER HI-BRED INT INC.
 PA
 XX Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
 PI WPI; 2001-514772/56.
 DR
 XX
 XX A male tissue-preferred regulatory region comprising nucleotide sequences
 PT essential for initiating transcription of the Ms45 gene useful for
 PT mediating fertility in a male plant.
 XX
 PS Example 5; Fig 8; 50bp; English.
 XX
 CC The invention provides a male tissue-preferred regulatory region (I)
 CC comprising nucleotide sequences essential for initiating transcription of
 CC the Ms45 gene. A method of mediating male fertility in a plant is
 CC provided that involves introducing an expression vector comprising a
 CC promoter operably linked to (I) into a plant where the exogenous gene
 CC impacts male fertility of the plant and (I) controls expression of the
 CC exogenous gene. A method of producing hybrid seeds is also provided. The
 CC present sequence represents a Z. mays Ms45 promoter fragment
 XX

SO Sequence 255 BP; 59 A; 86 C; 39 G; 71 T; 0 U; 0 Other;

Query Match 12.3%; Score 160.6; DB 5; Length 255;
 Best Local Similarity 97.6%; Pred. No. 1.3e-24;
 Matches 163; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1145 CGGCACTCTCGGTATATCTACATGCGATCTGTTTCAACCGTTGCTGTGT 1204
 DB 5 CCGCGATCCCGGTATATCTACATGCGATCTGTTTCAACCGTTGCTGTGT 64
 QY 1205 TCCATGTCGAAGCTTGGCTATTTGAAACCAAGAGATACCTACCTCCAAACATTCAT 1264
 DB 65 TCCATGTCGAAGCTTGGCTATTTGAAACCAAGAGATACCTACCTCCAAACATTCAT 124
 QY 1265 CTTACTCATGCAACTTCATGCAACGCGACATATGTTCCGAAC 1311
 DB 125 CTTACTCATGCAACTTCATGCAACGCGACATATGTTCCGAAC 171

RESULT 6
 AAH76334
 ID AAH76334 standard; DNA; 158 BP.
 XX
 AC AAH76334;
 XX

DT 29-OCT-2001 (first entry)
XX
DE 2. mays Ms45 male tissue-preferred regulatory region fragment.
XX
KW Ms45; male tissue; regulatory region; transcription; male fertility;
XX hybrid seed; db.
OS Zea mays.
XX
XX WO200160997-A2.
PN 23-AUG-2001.
XX
XX 13-FEB-2001; 2001WO-US004527.
PF
XX 15-FEB-2000; 2000US-00504487.
PR
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
PI WPI; 2001-514772/56.
XX
XX A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the Ms45 gene useful for
PT mediating fertility in a male plant.
XX
XX Claim 5; Page 47; 50pp; English.
XX
XX The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the Ms45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a DNA fragment -38 to -195 bases upstream of
CC the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
CC nucleotide sequence
XX
SQ Sequence 158 BP; 41 A; 50 C; 21 G; 46 T; 0 U; 0 Other;
Query Match 11.1%; Score 146; DB 5; Length 158;
Best Local Similarity 99.4%; Pred. No. 1.5e-21;
Matches 157; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1155 CGTGCATCTCACATGCGATCTACTACATGCTTGTTCACCGTTCGTC-TTGTTCATCGTC 1213
DB 1 CGTGCATCTCACATGCGATCTACTACATGCTTGTTCACCGTTCGTCCTTGTTCATCGTC 60
QY 1214 CAAGCCTTGCCCTTCTTCTGAACCAAGAGATCTACTCTCCCAACATCATCTTACTCAT 1273
DB 61 CAAGCCTTGCCCTTCTTGAACCAAGAGATCTACTCTCCCAACATCATCTTACTCAT 120
QY 1274 GCAACTTCCATGCAACACGCAATATGTTTCTGTAAC 1311
DB 121 GCAACTTCCATGCAACACGCAATATGTTTCTGTAAC 158

RESULT 7
AAL15210/c
ID AAL15210 standard; cDNA; 883 BP.
XX
AC AAL15210;
XX
DT 07-DEC-2001 (first entry)
XX
XX Human breast cancer expressed polynucleotide 7667.
DE Human breast cancer; cell marker; cytosolic; ss.
XX
XX Homo sapiens.
XX

PN WO200151628-A2.
XX
XX 19-JUL-2001.
PD
XX
XX 10-JAN-2001; 2001WO-US000798.
PF
XX
XX 14-JAN-2000; 2000US-0176077P.
PR 14-MAR-2000; 2000US-0189167P.
PR 24-MAR-2000; 2000US-0192099P.
PR 29-MAR-2000; 2000US-0193480P.
PR 15-MAY-2000; 2000US-0205230P.
PR 09-JUN-2000; 2000US-0211315P.
PR 25-JUL-2000; 2000US-0220534P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Lillie J, Xu Y, Wang Y, Steinmann K;
PI WPI; 2001-451856/48.
XX
XX
XX New peptide useful as a marker for the diagnosis of breast cancer.
PT
XX
XX Claim 1; Page 1378; 3695pp; English.
XX
XX The invention relates to human breast cancer expressed polynucleotides
CC (Aa107544-Aa126789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterizing treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity
XX
SQ Sequence 883 BP; 322 A; 18 C; 23 G; 396 T; 0 U; 124 Other;
Query Match 5.4%; Score 70.2; DB 4; Length 883;
Best Local Similarity 39.2%; Pred. No. 3.4e-05;
Matches 237; Conservative 0; Mismatches 367; Indels 0; Gaps 0;
QY 419 TGTCCACTATTTGGCACTACCCAGAACAGATTTAAAAAATTAACCAAGTAATCA 478
DB 873 TGTNNANNNAACCTAACCCTTTAAANATNTNNANTNNAANNAAAAAATTTAAANN 814
QY 479 CTCGAAGACTATCATGTATGTTTAAAGAAACATCTATTAACACGATCTCTTAAA 538
DB 813 TAAAAATTTTNNAAAACTAATTTNATNNTAATNTTNTTAAAAAANNNAAAAA 754
QY 539 AACAGCATATTTTGAAGAAGCAAAATTAATGTTACAGTTTCAACATCTAAGACGACA 598
DB 753 TTAANNNTTNTNATNTATTAACCAAAATTTTAAAAAATTTTAAAAAATTTT 694
QY 599 AATTATATCGAAAGGTAAGCTATGACGTTCAATTTTCTTTTCATTCTGTTAATTTG 658
DB 653 AATATNTTAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTT 634
QY 659 TTATGTTTATATATACATTTTCTCTCTTCAATAGAGTATTTCTCGATTTTATA 718
DB 633 TTTNTTNNATTAATAAAAAATTTTAAATTAATAAAAAAANNNTTTTAAANATTTTAA 574
QY 719 AAATGACTATAAGCATTTTATATATAAGACGACATGTCGTAGATTCGTCAAAAA 778
DB 573 AAAAAAANNTTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTT 514
QY 779 TCTTTCTGATTTTAAAAAGCTAGTTTGGCAACCGTGTCTTTCAAGAAATTTGAT 838
DB 513 AAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTNNNT 454
QY 839 TTTTCAAAAAAATTTAGTTTATTTCTCTTATATAAAGAAAAACATTAAGAAAAATAGA 898
DB 453 NNANATTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAATAT 394
QY 899 GTTGCAGACTAGCCCTAGAAATGTTTCCCAATAAATTAACAATCACTGTGATATATT 958

```

Db      393 TTTNAAAAAATTAATTTATATNTATNTNAAANAAATTTTAAATNNAANAAAAATTTTAA 334
Qy      959 TGGCCAGCCCCATTAATTTATTTAAACGAAATCGAGCAAAACCAATCTGAGCT 1018
Db      333 NTTAATAATAATAATAATTAATTTAAACATCATTTTNTAATTTTAAATTAATAAAAAAAT 274
Qy      1019 ATTT 1022
Db      273 NTTT 270

RESULT 8
ACN85231/C
ID      ACN85231 standard; DNA; 960 BP.
XX
XX      ACN85231;
XX
XX      02-DEC-2004 (first entry)
XX
XX      Breast cancer related marker. seq id 6381.
XX
XX      Cancer; breast; tumour; cytostatic; marker; detection; therapy; ds.
XX
XX      Homo sapiens.
XX
XX      US2003099974-A1.
XX
XX      29-MAY-2003.
XX
XX      18-JUL-2002; 2002US-00198846.
XX
XX      18-JUL-2001; 2001US-0306220P.
XX
XX      (MILL-) MILLENNIUM PHARM INC.
XX
XX      Lillie J, Xu Y, Wang Y, Steinmann K;
XX
XX      WPI; 2003-787014/74.
XX
XX      Novel isolated polypeptide associated with breast cancer, useful for
XX      detecting presence of polypeptide in sample, as a marker for breast
XX      cancer.
XX
XX      Disclosure; SEQ ID NO 6381; 36bp; English.
XX
XX      The invention relates to an isolated polypeptide (I) associated with
XX      breast cancer which is encoded by a nucleic acid molecule comprising a
XX      nucleotide sequence (SI). Further disclosed is an antibody that binds to
XX      the polypeptide of the invention. The activity of the polypeptide of the
XX      invention may be described as cytostatic. The antibody is useful for
XX      detecting the presence of (I) in a sample. Nucleic acid molecules of the
XX      invention are useful in the detection of breast tumors. (I) is useful as
XX      a marker for breast cancer and in breast cancer therapy. Sequences given
XX      in records ACN78851-ACN92934 represent nucleic acid markers associated
XX      with breast cancer. Note: The sequence listing does not form part of the
XX      specification but may be obtained in electronic format from the USPTO web
XX      site at seqdata.uspto.gov/sequence.html?DocID=20030099974
XX
XX      Sequence 960 BP; 340 A; 33 C; 39 G; 421 T; 0 U; 127 Other;
XX
Query Match      5.4%; Score 70.2; DB 11; Length 960;
Best Local Similarity 39.2%; Pred. No. 3.5e-05;
Matches 237; Conservative 0; Mismatches 367; Indels 0; Gaps 0;
Qy      419 TGTCCACTATGCGACCTAGCAGACAGATTTAAAAAATAACCAAGTAATATCCCA 478
Db      933 TGTNNNNNNNANACCTTAACCTTTAAANATNTNNNATTTNANNAAAAAATATTTAANN 874
Qy      479 CTCGAAAGCTATCATGTATGTTTAAAGAAACATCTATTAAACACGATCCTTTAAA 538
Db      873 TAAATAATTTTAAATAAACTAATTTNNAANNATNTAATTTTNTTAAAAAANNAAAAA 814

```

```

Qy      539 AACGACATATTTGGAAGAGCAATTAATGTTACAGTTTACAAACATCTAAGAGGACA 598
Db      813 TTAANNNTTTTNTAANTTAATTAACAACAAATTTTTTTTTTAAAAAATTTTAAANNT 754
Qy      599 AATTATATCGAAAGTAAGCTATGACGTTTCAGATTTTCTGTTTCATCTGTTATTTTG 658
Db      753 AATTAATTAATTTTNTTAATTAATAAAAAAATTTTNAATTTTAANCANNTNTTTT 694
Qy      659 TTATGTTTAAATATACATTTCTCTCTTACATAGAGTATTTTCTCCGATTTTATA 718
Db      693 TTTNTTTNNATATAAAAAAATTTTAAATATAAAAAAANNNTTTTATNAAANATTTAA 634
Qy      719 AATGACTATTAAGTATTTTATATAAGACGAGATGTCGAGATCTCGTCAAAA 778
Db      633 AAAAAAANNNTTTTTTTTATANTTTATAAAAAATTTATNTNTTTTTCTNAAAAANA 574
Qy      779 TCTTCTGATTTTTTTTAAAGCTAGTTTGGCAACCTGTTCTTCAAGAATTTGATT 838
Db      573 AAAAAAATTTTNTNAAANNTTTTAAAAACCTTAATTAANNAANAAATTTTTNNNTT 514
Qy      839 TTTTCAAAAAAATTAATTTATTTTCTCTTATTAATAAGAAAAACATTAGAAAAATAGA 898
Db      513 NNNATTTAAATAATTTATTTTNTCTAATTAATAAANNAANNAATTAATTAAT 454
Qy      899 GTTCCAGACTAGCCTTAGAATGTTTCCCAATTAATTAATCACTGCTATTAATT 958
Db      453 TTTNAAAAAATTAATTTATANTTAATNNAANNAATTTAAANTNANNAAAAAATTTTA 394
Qy      959 TGGCCAGCCCCATTAATTTATTTAAACGAAATCGAAATGAGAGAAACCAATCTGAGCT 1018
Db      393 NTTAATAATAATAATAATTAATTTAAACATCATTTTNTAATTTAATTAATAAAAAAAT 334
Qy      1019 ATTT 1022
Db      333 NTTT 330

```

```

RESULT 9
ADR04296
ID      ADR04296 standard; DNA; 13400 BP.
XX
XX      ADR04296;
XX
XX      04-NOV-2004 (first entry)
XX
XX      Corn FT homologue nucleotide sequence SEQ ID NO:63.
XX
XX      flowering locus T gene; FT; terminal flower; TFL; Apetala3; Ap3; plant;
XX      floral development; plant sterility; plant fertility; flowering time;
XX      plant growth rate; inflorescence architecture; tissue culture morphology;
XX      cell division; FT homologue; gene; ds.
XX
XX      Zea mays.
XX
XX      PN MO2004067723-A2.
XX
XX      12-AUG-2004.
XX
XX      29-JAN-2004; 2004WO-US002422.
XX
XX      30-JAN-2003; 2003US-00343477.
XX
XX      (PION-) PIONEER HI-BRED INT INC.
XX      (DUPO ) DU PONT DE NEMOURS & CO B I.
XX
XX      Danilevskaya O, Hermon P, Bruggemann E, Shitroun D, Ananiev E;
XX      Rafalski JA, Sakai H, Cahoon E, Cahoon R, Klein T;
XX
XX      WPI; 2004-580996/56.
XX
XX      New polymucleotides, specifically nucleic acid fragments encoding
XX      flowering locus T gene (FT) or terminal flower (TFL), or Apetala3 (Ap3)
XX      homologs, useful for floral development, e.g. engineering plant flowering

```


FT	/note= "functional fragment of cDNA specifically claimed in Claim 16"
FT	120. .5147
FT	/*tag= a
FT	120. .1221
FT	/*tag= b
FT	/note= "functional fragment of cdna specifically claimed in Claim 12"
FT	565. .816
FT	/*tag= d
FT	/note= "functional fragment of cdna specifically claimed in Claim 14"
FT	655. .1221
FT	/*tag= c
FT	/note= "functional fragment of cdna specifically claimed in Claim 15"
FT	1438. .2424
FT	/*tag= g
FT	/note= "functional fragment of cdna specifically claimed in Claim 17"
FT	2425. .3791
FT	/*tag= h
FT	/note= "functional fragment of cdna specifically claimed in Claim 18"
PN	NO9924575-A1.
PD	20-MAY-1999.
PF	12-NOV-1998; 98MO-US024225.
PX	12-NOV-1997; 97US-00968542.
PA	(IOWA) UNIV IOWA STATE RES FOUND INC.
PI	Myers AM, James MG;
PP	WPI, 1999-327406/27.
DR	P-PDSB; AAY06199.
XN	Nucleic acid encoding starch synthase enzyme of maize.
XX	Claim 1; Page 104-107, 138pp; English.
PS	This is the nucleotide sequence of the maize gene dui1 (du1). To
XX	illustrate the role of the du1 locus in starch biosynthesis, a transposon
CC	tagging strategy was used to isolate the gene and describe its
CC	polypeptide product. The invention reports tagging of the du1 locus with
CC	Mu transposon, cloning and characterization of a portion of the gene, and
CC	isolation a near full-length cDNA (the present sequence). The amino acid
CC	sequence (see AAY06199) deduced from this cDNA indicates that Dui codes
CC	for a 186 kDa protein extremely similar to potato tuber starch synthase
CC	SSIII. Its expression pattern indicates that Dui codes for SSII of maize
CC	cDopernm. The Dui product contains unique sequence features in its N-
CC	terminus that may mediate direct interactions with other starch
CC	biosynthetic enzymes. Mutations within the maize SSII gene affect
CC	multiple aspects of starch biosynthesis by disrupting an enzyme complex
CC	containing starch synthase(s), starch branching enzyme(s) and possibly an
CC	starch debranching enzyme(s). The isolated cDNA can be used to provide an
CC	enzyme with which to regulate the production of starch, and with which to
CC	produce altered or novel forms of starch, e.g. in transgenic plants.
CC	Expression of Dui in bacteria and yeasts also modifies glycogen
CC	production. Claimed expression vectors comprise the cDNA or fragments of
CC	it that code for functional portions of Dui
XX	Sequence 6027 BP; 1800 A; 1177 C; 1441 G; 1609 T; 0 U; 0 Other;
Query Match	5.1%; Score 66.4; DB 2; Length 6027;
Best Local Similarity	77.5%; Pred No. 0.00035;

	Matches	93;	Conservative	0;	Mismatches	26;	Indels	1;	Gaps	1;
Oy	792	TTTAAAGACTGATGGCAACCCGTCTTCTTCATAAAGAATTGATTTTTTCAAAAAAA	851							
Dd	5606	TCTACGGGCTAGTTGGGAACCACATT-TTCCAAGGATTTCCATTTTCCAGAAAA	5548							
Oy	852	TTAGTTATTTTCTCTTTATTAATAATAGAAAAACCTTAGAAAAATAGAGTTGCCAGCTAG	911							
Dd	5547	TTAGTTATTTTTCATTGGAAAAATTAATAATCTCTTGAAAAAATAGAGTTACATATCTAG	5488							
RESULT 12										
ID	ABX09935/C									
XX	ABX09935	standard; DNA; 6027 BP.								
XX	AC									
XX	ABX09935;									
DT	17-FEB-2003	(first entry)								
XX	DE	DNA encoding maize Starch synthase III (Dui).								
XX	KM	Starch; starch synthase; glucan association domain; GLASS; linker domain,								
XX	KM	LINKR; glucosyl transferase domain; GXYTR; C-terminal end; CTEND;								
XX	KW	granule bound starch synthase; GBSS; morphology; retrogradation;								
XX	OS	waterbinding; swelling potential; gene; ds.								
XX	Zea mays.									
PM	WO200279410-A2.									
PD	10-OCT-2002.									
PF	29-MAR-2002; 2002WO-US009574.									
PR	30-MAR-2001; 2001US-0279720P.									
PA	(BADI) BASF PLANT SCI GMBH.									
PI	Commuri P, Keeling PL, Ramirez N, McKean A, Gao Z, Guan H;									
XX	WP1; 2003-040678/03.									
PT	New DNA encoding fusion protein consisting of 4 different functional									
PT	domains selected from glucan association domain, linker domain, glucosyl									
PT	transferase domain, and C-terminal end, useful for producing modified									
PT	starches.									
PS	Claim 33; Page 225-227; 265pp; English.									
XX	The invention describes an isolated DNA molecule encoding a fusion									
XX	protein consisting of 4 different functional domains selected from glucan									
XX	association domain (GLASS), linker domain (LINKR), glucosyl transferase									
XX	domain (GXYTR), and C-terminal end (CTEND) which are operably linked to									
XX	one another. The DNA molecule is useful for expressing in plants									
XX	polypeptides including starch synthase enzymes as fusion proteins with									
XX	improved affinity to starch and modified catalytic capabilities and to									
XX	the in vivo and in vitro synthesis of glucan chains of modified lengths									
XX	as compared to plants producing native starch or starch produced with									
XX	native starch synthases. Expression of the starch synthase fusion									
XX	proteins along with granule bound starch synthase (GBSS) will lead to a									
XX	modified starch having an altered or improved morphology, retrogradation,									
XX	waterbinding, or swelling potential of the granules, gel strength,									
XX	adhesiveness, cohesiveness, hardness, elasticity, increased or decreased									
XX	granule size, degree of branching, crystallinity, degree of cross-									
XX	linking, and increased or decreased glucan chain lengths. This sequence									
XX	encodes a starch synthase used in the invention									
SQ	Sequence 6027 BP; 1800 A; 1177 C; 1441 G; 1609 T; 0 U; 0 Other;									
Query Match	5.1%; Score 66.4; DB 8; Length 6027;									
Best Local Similarity	77.5%; Pred. No. 0.00035;									
Matches	93; Conservative 0; Mismatches 26; Indels 1; Gaps 1;									

QY	792	TTTAAAGACTGTTTGGGAAACCCGTTCTCTTCAAGAAATTGATTTTTCACAAAAAAA	851
QY	792	TTTAAAGACTGTTTGGGAAACCCGTTCTCTTCAAGAAATTGATTTTTCACAAAAAAA	851
Db	5606	TCCTAGGGCTAGTTGGGAAACCCCATTT-TTCCAGAGGATTTCCATTTTTCAGAAAAAA	5548
QY	852	TTAGTTTATTTCTCTTATATAAATAGAAAAACATTAGAAAAATAGAGTTGCCAGACTAG	911
Db	5547	TTAGTTTATTTCTCTTATATAAATAGAAAAACATTAGAAAAATAGAGTTGCCAGACTAG	5488

RESULT 13

ID	ADK12106 standard; cDNA, 6027 BP.
XX	ADK12106;
AC	
XX	03-JUN-2004 (first entry)
DT	
XX	cDNA encoding maize starch synthase, DUL.
DE	
XX	Maize, dul1 1; DUL; starch synthase;
KW	alpha-1,4-glycosyltransferase catalytic activity; starch production;
KV	plant; gene; ss.
XX	
OS	Zea mays.
XX	
FH	Key
FT	Location/Qualifiers
FT	CDS
FT	120..514
FT	/tag= a
XX	/product= "DUL"
XX	
PN	US2004049810-A1.
XX	
PD	11-MAR-2004.
XX	
PF	05-AUG-2003; 2003US-00634262.
XX	
PR	12-NOV-1997; 97US-00968467.
PR	12-MAY-2000; 2000US-00554467.
XX	
PA	(MYER/) MYERS A M.
PA	(JAME/) JAMES M G.
XX	
PI	Myers AM, James MG;
XX	
DR	WPI; 2004-238526/22.
XX	
DR	P-PSDB; ADK12117.
XX	
PT	New nucleic acid designated dul1, encoding a starch synthase, useful in
PT	starch production.
XX	
PS	Claim 1; SEQ ID NO 1; 58pp; English.
XX	
CC	The present invention relates to the isolation of a maize gene, dul1 1
CC	(DUL), and the polypeptide it encodes. The DUL polypeptide has starch
CC	synthase activity, and comprises an N-terminal arm region, a C-terminal
CC	catalytic region, and a region of about 900 amino acids terminating with
CC	the catalytic region. The C-terminal catalytic region has a catalytic
CC	domain comprising alpha-1,4-glycosyltransferase catalytic activity. The
CC	dul polynucleotide sequence is useful in producing starch e.g. from a
CC	transgenic plant or transfected cell. The present sequence encodes maize
CC	DUL.
XX	
SO	Sequence 6027 BP; 1800 A; 1177 C; 1441 G; 1609 T; 0 U; 0 Other;
XX	
Query Match	5.1%; Score 66.4; DB 12; Length 6027;
Best Local Similarity	77.5%; Pred. No. 0.00035;
Matches	93; Conservative 0; Mismatches 26; Indels 1; Gaps 1;
OY	792 TTTAAGAGCTAGTTGGCAACCGTCTTCTTCAAGATTTGATTTTCAAAAAA 851
Db	5606 TCTACGGGCTAGTTGGAGACCCCATTT-TTCCAAGGATTTCCATTTTCAAGAAAA 5548
OY	852 TTACATTTATTTCTTTATATAATAGAAAAACACTTAGAAAAATAGAGTTGCCAGACTAG 911

5547 TTAGTTAATTTTCATGGAAAAATTGAATCTCTTGGAAAAATAGACTTCACATACTAG 5488

RESULT 14

ABX35844	ABX35844 standard; cDNA; 439 BP.
ID	ABX35844
XX	
XX	
AC	ABX35844;
XX	
DT	20-FEB-2003 (first entry)
XX	
XX	
DE	Bovine EST associated with lactation/muscle/fat deposition #1009.
XX	
XX	
KM	Bovine; aa; EST; expressed sequence tag; lactation; LMFD;
KM	muscle deposition; fat deposition; genome mapping; gene identification;
KM	gene analysis; cattle breeding.
XX	
XX	
DS	Bos Taurus.

PN US2002137139-A1

AA		
PD	26-SEP-2002.	
XX		
XX		
PF	24-SEP-2001; 2001US-00960352	
XX		
PR	12-JAN-1999; 99US-0115707P.	
PR	11-JAN-2000; 2000US-00480902.	

PA (BYAT/) BYATT J C.

PA (MATH/) MATHIALAGAN N
PA (TAON/) TAO N.
PA (WARR/) WARREN W C.

PI Byatt JC, Mathialagan N, Tao N, Warren WC;

DR WPI; 2003-110599/10.

PT New nucleic acid associated with lactation, and muscle and fat
PT deposition, useful for genome mapping, gene identification and analysis,
PT cattle breeding, or for genetically improving cattle.

PS Claim 2; SEQ ID NO 1009; 245bp; English.

The invention relates to a purified nucleic acid molecule associated with lactation or muscle and fat deposition (designated LMPD), derived from cattle, and the LMPD nucleic acid can specifically hybridize to a second nucleic acid molecule comprising any of 15112 nucleotide sequences, appearing as ABX3483c-ABX49947, or complements of them. Also included are: (1) a transformed cell having a nucleic acid comprising an LMPD nucleic acid linked to a promoter and a 3' non-translated sequence that functions in the cell to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and (2) determining a level or pattern of a molecule in a bovine cell or tissue comprising: (a) incubating a marker nucleic acid (comprising any of the 15112 nucleic acid sequences or its complement or fragment) with a complementary nucleic acid molecule obtained from the bovine cell or tissue, where hybridisation between the marker nucleic acid and the complementary nucleic acid permits the detection of the molecule; and (b) detecting the level or pattern of the complementary nucleic acid, where the detection of the complementary nucleic acid is predictive of the level or pattern of the molecule. The LMPD nucleic acid is used for determining a level or pattern of a molecule in a bovine cell or tissue. It is useful for genome mapping, gene identification and analysis, cattle breeding, preparation of constructs for use in cattle gene expression, or for genetically improving cattle. The present sequence is one of the 15112 bovine LMPD EST (expressed sequence tag) nucleic acids. Note: The present sequence was not shown in the specification but was obtained in electronic format from the USPTO web site:

CC seqdata.uspto.gov/sequence.html?DocID=20020137139
XX
SQ Sequence 439 BP; 45 A; 51 C; 56 G; 286 T; 0 U; 1 Other;

Sequence 439 BP; 45 A; 51 C; 56 G; 286 T; 0 U; 1 Other;

Query Match 5.0%; Score 66.2; DB 8; Length 439;
 Best Local Similarity 54.3%; Pred. No. 0.0002; Mismatches 113; Indels 0; Gaps 0;
 Matches 134; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

QY 626 TTCAGATTTTCTTTTCATTCGTGTAATTTGTAATGTTTAAATACATTTTCTCT 685
 DB 182 TTCGATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 241
 QY 686 CTTACATAGAGTATTTCTCCGATTTTAAATAGCATTAAGTCAATTTTATATA 745
 DB 242 TTTTATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 301
 QY 746 AGACGACGATGCTGATTCGTCGTAATTCCTGATTTTCTGATTTTCTGATTTTCT 805
 DB 302 TGTGGGGGGTGTGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 361
 QY 806 TGGCAACCTGTTCTTTCTTCAAGAAATTTGATTTTCTCAAAAAAATAGTTATTTCT 865
 DB 362 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 421
 QY 866 CTTTATA 872
 DB 422 TTTCTTA 428

RESULT 15
 AA187279/C
 ID AA187279 standard; cDNA; 346 BP.
 XX
 AC AA187279;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 7339.
 XX
 DE Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KM vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KM tissue growth factor; immunomodulatory; cancer; leukaemia;
 KM nervous system disorders; arthritis; inflammation; ss.
 XX
 OS Homo sapiens.
 XX
 PN MO200164835-A2.
 PD 07-SEP-2001.
 PF 26-FEB-2001; 2001WO-US004927.
 PR 28-FEB-2000; 2000US-00515126.
 PR 18-MAY-2000; 2000US-00577409.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 DR WPI; 2001-514838/56.
 DR P-PDB; AA007348.
 XX
 PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing
 PT and treating e.g. leukemia, inflammation and immune disorders.
 XX
 PS Claim 1; SEQ ID NO 7339; 1399bp + Sequence Listing; English.

CC The invention relates to human polynucleotides (AA179941-AA193841) and
 CC the encoded proteins (AA000010-AA013910) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and

CC inflammation. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX

SQ Sequence 346 BP; 188 A; 22 C; 16 G; 120 T; 0 U; 0 Other;
 Query Match 5.0%; Score 65.6; DB 4; Length 346;
 Best Local Similarity 52.6%; Pred. No. 0.00026; Mismatches 129; Indels 0; Gaps 0;
 Matches 143; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

QY 446 AGATTAATAAATAAACAAGTAACTAAATCCATCGAAGATATCATGTAATGTTTAA 505
 DB 334 ACAATTAAGATATATATTAATTAATTAAGATTAATTAATTAATTAATTAATTA 275
 QY 506 GAAACATCTATTAAACCAAGATCTTTAAATAACAAGCATTTTGAAGAAGACAAAT 565
 DB 274 GACTTATATATATATTAATTAATTTTCTTTAGAAATTAATTAATTAATTAATTA 215
 QY 566 TATGTTACGTTTACAAACATCTAAGAGCAAAATTAATGAAAGTAAAGTATGAG 625
 DB 214 TTTTAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 155
 QY 626 TTCAGATTTTCTTTTCATTCGTGTAATTTGTAATGTTTAAATACATTTTCTCT 685
 DB 154 TATTTATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 95
 QY 686 CTTACATAGAGTATTTCTCCGATTTTAT 717
 DB 94 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 63

Search completed: September 15, 2005, 21:33:20
 Job time : 635.708 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 14, 2005, 19:14:19 ; Search time 198.304 Seconds
(without alignments)
10817.505 Million cell updates/sec

Title: US-10-713-381-2_COPY_1_1311

Perfect score: 1311
Sequence: 1 ccatggcgcctctatcgaaaa.....cgcacatcgtctcgaac 1311

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B COMB.seq: *
3: /cgn2_6/prodata/1/ina/6A COMB.seq: *
4: /cgn2_6/prodata/1/ina/6B COMB.seq: *
5: /cgn2_6/prodata/1/ina/6C COMB.seq: *
6: /cgn2_6/prodata/1/ina/6D COMB.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1311	100.0	1394	3	US-08-880-499-1
2	1311	100.0	1394	3	US-08-880-499-2
3	66.4	5.1	6027	2	US-08-968-542C-1
4	66.4	5.1	6027	4	US-09-554-467A-1
5	58.6	4.5	2523	2	US-08-410-784A-3
6	57	4.3	19124	2	US-08-487-826B-13
7	55.8	4.3	279	4	US-09-313-294A-5397
8	55.6	4.2	612	4	US-09-302-540-1357
9	55.6	4.2	1394	3	US-08-880-499-1
10	55.6	4.2	1394	3	US-08-880-499-2
11	55.4	4.2	1055	4	US-09-806-708B-23
12	55.2	4.2	2614	4	US-09-004-056-1
13	55	4.2	396	4	US-09-640-173-53
14	55	4.2	396	4	US-09-713-550-53
15	55	4.2	396	4	US-09-825-294-53
16	55	4.2	396	4	US-09-970-866-53
17	54.6	4.2	55886	4	US-09-949-016-15129
18	53.8	4.1	307	4	US-09-313-294A-4743
19	53.8	4.1	6027	2	US-08-968-542C-1
20	53.8	4.1	6027	4	US-09-554-467A-1
21	52.2	4.0	1141	4	US-09-806-708B-22
22	51.6	3.9	134987	4	US-09-949-016-15348
23	51.6	3.9	134987	4	US-09-949-016-15349
24	51.6	3.9	134987	4	US-09-949-016-15350
25	51.6	3.9	134987	4	US-09-949-016-15507
26	51.6	3.9	134987	4	US-09-949-016-15508
27	51.6	3.9	134987	4	US-09-949-016-15509

C	28	51.4	3.9	50383	4	US-09-949-016-17600	Sequence 17600, A
C	29	51.4	3.9	129415	4	US-09-949-016-16997	Sequence 16997, A
C	30	51	3.9	601	4	US-09-949-016-156535	Sequence 156535, A
C	31	51	3.9	1141	4	US-09-806-708B-22	Sequence 22, Appl
C	32	50.6	3.9	1039	4	US-09-902-540-1280	Sequence 1280, Ap
C	33	50.6	3.9	16573	4	US-09-949-016-14876	Sequence 14876, A
C	34	50.6	3.9	18773	4	US-09-949-016-14164	Sequence 14164, A
C	35	50.4	3.8	731	1	US-08-451-405A-2	Sequence 2, Appl
C	36	50.2	3.8	832	4	US-09-621-976-2813	Sequence 2813, Ap
C	37	50.2	3.8	317366	4	US-09-949-016-16001	Sequence 16001, A
C	38	50	3.8	1392	3	US-09-257-584-1	Sequence 1, Appl
C	39	49.8	3.8	2435	3	US-09-306-593-1	Sequence 1, Appl
C	40	49.8	3.8	231129	4	US-09-949-016-16110	Sequence 16110, A
C	41	49.8	3.8	266293	4	US-09-949-016-11934	Sequence 11934, A
C	42	49.6	3.8	640681	4	US-09-790-988-1	Sequence 1, Appl
C	43	49.2	3.8	126176	4	US-09-949-016-16137	Sequence 16137, A
C	44	49.2	3.8	126176	4	US-09-949-016-16138	Sequence 16138, A
C	45	47.8	3.6	601	4	US-09-949-016-25787	Sequence 25787, A

ALIGNMENTS

RESULT 1
US-08-880-499-1
; Sequence 1, Application US/08880499
; Patent No. 6037523
; GENERAL INFORMATION:
; APPLICANT: Albertson, Marc C.
; APPLICANT: Fox, Tim W.
; APPLICANT: Carl, Garnaat W.
; APPLICANT: Hultman, Gary A.
; APPLICANT: Kendall, Timmy L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION
; TITLE OF INVENTION: AND METHOD OF USING SAME
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
; STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.
; STREET: Box 1000
; CITY: Johnston
; STATE: Iowa
; COUNTRY: USA
; ZIP: 50131
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/880,499
; FILING DATE: CONCURRENTLY HEREWITH
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Sweeney, Patricia A.
; REGISTRATION NUMBER: 32,733
; REFERENCE/DOCKET NUMBER: 0578
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (515) 248-4800
; TELEFAX: (515) 248-4844
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1394 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-880-499-1

Query Match 100.0%; Score 1311; DB 3; Length 1394;
Best Local Similarity 100.0%; Pred. No. 1.1e-313;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCATGAGTCTCTATGAAAAAGATGAGTACATGTGTCTATATCCGTTTCTTAGGGTCC 60
DB 1 CCATGAGTCTCTATGAAAAAGATGAGTACATGTGTCTATATCCGTTTCTTAGGGTCC 60
QY 61 CTTCTTCTGCTTATTAAGTAAATCGGGGTTTACAAAACTTCCAGGGTGCATGAT 120
DB 61 CTTCTTCTGCTTATTAAGTAAATCGGGGTTTACAAAACTTCCAGGGTGCATGAT 120
QY 121 CTCGATGTTCACTTCTCCCACTCGCGTTCACATTTCTTGGATGCGGGTGTCCCAT 180
DB 121 CTCGATGTTCACTTCTCCCACTCGCGTTCACATTTCTTGGATGCGGGTGTCCCAT 180
QY 181 CTGACCGAGGCCCATGACACCTTTGCGAGACCCCATGAGGCGCTTTCGATGCGCA 240
DB 181 CTGACCGAGGCCCATGACACCTTTGCGAGACCCCATGAGGCGCTTTCGATGCGCA 240
QY 241 CGAGAGCTATCGGGTGTGTGTATCCAGGGGATATATGTCTCCCAATGCTACCTATA 300
DB 241 CGAGAGCTATCGGGTGTGTGTATCCAGGGGATATATGTCTCCCAATGCTACCTATA 300
QY 301 TTATTAATCTTATGATATTAATTAATTTTGAATAAACAATTAATCTTCTTGTGTA 360
DB 301 TTATTAATCTTATGATATTAATTAATTTTGAATAAACAATTAATCTTCTTGTGTA 360
QY 361 GGGCTCAGCAATGATTTTCTGTTAGGGCCAGAAAATGCGAGGACCGCATGCTAGTG 420
DB 361 GGGCTCAGCAATGATTTTCTGTTAGGGCCAGAAAATGCGAGGACCGCATGCTAGTG 420
QY 421 TCCACTATTTGGCACTCCAGAAACAGATTTTAAATAATTAACAAAGTAACTATCACT 480
DB 421 TCCACTATTTGGCACTCCAGAAACAGATTTTAAATAATTAACAAAGTAACTATCACT 480
QY 481 CGAAGCTATCATGATGATTTTAAAGAAACATCTATTAACACGAGCATCTTAAATAA 540
DB 481 CGAAGCTATCATGATGATTTTAAAGAAACATCTATTAACACGAGCATCTTAAATAA 540
QY 541 CAAGCATATTTGCAAAAGACAAATTAATGTTTACAGTTTACAAATCTAAGACGCAAA 600
DB 541 CAAGCATATTTGCAAAAGACAAATTAATGTTTACAGTTTACAAATCTAAGACGCAAA 600
QY 601 TTATATCGAAAGTATGATGATGCTGATGATTTTCTTTCATCTTGTATTTGTT 660
DB 601 TTATATCGAAAGTATGATGATGCTGATGATTTTCTTTCATCTTGTATTTGTT 660
QY 661 ATTGTTTATATACATTTTCTCTCTTCAATAGAGTATTTTCCGATTTTAA 720
DB 661 ATTGTTTATATACATTTTCTCTCTTCAATAGAGTATTTTCCGATTTTAA 720
QY 721 ATGACTATTAAGTATTTTATTAAGACGCGATGCTGATGATTTCTGTTCAAAATC 780
DB 721 ATGACTATTAAGTATTTTATTAAGACGCGATGCTGATGATTTCTGTTCAAAATC 780
QY 781 TTTCGATTTTAAAGCTAGTTTGGCAACCCGTTTCTTCAAGAAATTTGATTTT 840
DB 781 TTTCGATTTTAAAGCTAGTTTGGCAACCCGTTTCTTCAAGAAATTTGATTTT 840
QY 841 TTCAAAAAAATAGTTTATTTTCTCTTAAATAAGAAAACTTAGAAAAATAGAGT 900
DB 841 TTCAAAAAAATAGTTTATTTTCTCTTAAATAAGAAAACTTAGAAAAATAGAGT 900
QY 901 TGGCAGACTAGCCCTAGAAATGTTTCCCAATAAATTAACAATCACTGTGATATTAATTTG 960
DB 901 TGGCAGACTAGCCCTAGAAATGTTTCCCAATAAATTAACAATCACTGTGATATTAATTTG 960
QY 961 GCGAGCCCATTAATTAATTAACCGAAATCGAAATCGGAAACCAATCTGAGCTAT 1020
DB 961 GCGAGCCCATTAATTAATTAACCGAAATCGAAATCGGAAACCAATCTGAGCTAT 1020
QY 1021 TTCTCTAGATTAGTAAAGAGAGAGAGAGAGAAATCAATTTAAGTATGTTGCC 1080
DB 1021 TTCTCTAGATTAGTAAAGAGAGAGAGAGAGAAATCAATTTAAGTATGTTGCC 1080
QY 1081 TGAGATGTGCGGTTTGGCAAGATAGCACCGTAAATCATAGTCAATAGTGTGCTAGTCA 1140

DB 1081 TGAGATGTGCGGTTTGGCAAGATAGCACCGTAAATCATAGTCAATAGTGTGCTAGTCA 1140
QY 1141 GGTTCGCGAGCTCTGTGTCTATCTCATGATGAGTACTATCTGTTTCAACCGTTGTC 1200
DB 1141 GGTTCGCGAGCTCTGTGTCTATCTCATGATGAGTACTATCTGTTTCAACCGTTGTC 1200
QY 1201 TTGTTCCATCTGTCGAAGCTTGCCTATTCTGAACCAAGAGATACCTACTCCCAACAT 1260
DB 1201 TTGTTCCATCTGTCGAAGCTTGCCTATTCTGAACCAAGAGATACCTACTCCCAACAT 1260
QY 1261 CCATCTTAATCTGCAATCTTCCATGCAACGCAACATATGTTTCTGAAC 1311
DB 1261 CCATCTTAATCTGCAATCTTCCATGCAACGCAACATATGTTTCTGAAC 1311

RESULT 2
US-08-880-499-2
Sequence 2, Application US/08880499
Patent No. 6037523
GENERAL INFORMATION:
APPLICANT: Albertson, Marc C.
APPLICANT: Fox, Tim W.
APPLICANT: Carl, Garnat W.
APPLICANT: Huffman, Gary A.
APPLICANT: Kendall, Jimmy L.
TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION
TITLE OR INVENTION: AND METHOD OF USING SAME
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.
CITY: Johnston
STATE: Iowa
COUNTRY: USA
ZIP: 50131
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/880,499
FILING DATE: CONCURRENTLY HEREWITH
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Sweeney, Patricia A.
REGISTRATION NUMBER: 32,733
REFERENCE/DOCKET NUMBER: 0578
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 248-4800
TELEFAX: (515) 248-4844
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1394 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-880-499-2

Query Match 100.0%; Score 1311; DB 3; Length 1394;
Best Local Similarity 100.0%; Pred. No. 1.1e-313;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCATGAGTCTCTATGAAAAAGATGAGTACATGTGTCTATATCCGTTTCTTAGGGTCC 60
DB 1 CCATGAGTCTCTATGAAAAAGATGAGTACATGTGTCTATATCCGTTTCTTAGGGTCC 60
QY 61 CTTCTTCTGCTTATTAAGTAAATCGGGGTTTACAAAACTTCCAGGGTGCATGAT 120
DB 61 CTTCTTCTGCTTATTAAGTAAATCGGGGTTTACAAAACTTCCAGGGTGCATGAT 120

```

QY 121 CTCGATGTTCCACTTCTCCCACTCGGTTGCACTTTCTTGATGTCGGTGGTCCCAT 180
DB 121 CTCGATGTTCCACTTCTCCCACTCGGTTGCACTTTCTTGATGTCGGTGGTCCCAT 180
QY 181 CTGACCGAGGCCCATGAGACACCTTTGAGGACACCATCAAGGGCTTTGATGGACCA 240
DB 181 CTGACCGAGGCCCATGAGACACCTTTGAGGACACCATCAAGGGCTTTGATGGACCA 240
QY 241 CGAGACGTATCGGGTCGTGGTGTATCCAGGGATATATGTCCTCCCAATGTCACCTATA 300
DB 241 CGAGACGTATCGGGTCGTGGTGTATCCAGGGATATATGTCCTCCCAATGTCACCTATA 300
QY 301 TTATTTCTTTGATATATATTTTATTTTGGAAAAATATACAACTTATCTTTGTGTA 360
DB 301 TTATTTCTTTGATATATATTTTATTTTGGAAAAATATACAACTTATCTTTGTGTA 360
QY 361 GGGCTCAGCATATGATTTTTCGCTTAAAGGCGCAGAAATGCGAGGACGACCTCTAGTG 420
DB 361 GGGCTCAGCATATGATTTTTCGCTTAAAGGCGCAGAAATGCGAGGACGACCTCTAGTG 420
QY 421 TCCACTATTGGCACTACCCAGAAACAAGTTTAAAAAATACCAAGTAACTATCCACT 480
DB 421 TCCACTATTGGCACTACCCAGAAACAAGTTTAAAAAATACCAAGTAACTATCCACT 480
QY 481 CGAAAGCTATCAGTATATGTTTAAAGAAACATCTATTTAAACACGATCTCTTAAAAA 540
DB 481 CGAAAGCTATCAGTATATGTTTAAAGAAACATCTATTTAAACACGATCTCTTAAAAA 540
QY 541 CAAGCATATTTGAAAAGAGACAAATATGTTACGTTTCAAAATCTTAAGAGCGACAA 600
DB 541 CAAGCATATTTGAAAAGAGACAAATATGTTACGTTTCAAAATCTTAAGAGCGACAA 600
QY 601 TTATATCGAAAGGTAAAGCTATGACGTCAGATTTTCTTTTCATCTTTATTTTGT 660
DB 601 TTATATCGAAAGGTAAAGCTATGACGTCAGATTTTCTTTTCATCTTTATTTTGT 660
QY 661 ATTGTTTTATATACATTTTCTTCTTACATAGATGATTTTCTTCGATTTTAA 720
DB 661 ATTGTTTTATATACATTTTCTTCTTACATAGATGATTTTCTTCGATTTTAA 720
QY 721 ATGACTATTAAGCATTTTATATTAAGAGACGATGCTAGATTTCTGTTCAAAATC 780
DB 721 ATGACTATTAAGCATTTTATATTAAGAGACGATGCTAGATTTCTGTTCAAAATC 780
QY 781 TTTCTGATTTTAAAGAGCTAGTTGGCAACCTGTTTCTTCAAGAAATTTGATTT 840
DB 781 TTTCTGATTTTAAAGAGCTAGTTGGCAACCTGTTTCTTCAAGAAATTTGATTT 840
QY 841 TTCAAAAAAAATTAAGTTATTTTCTTATTAAGAAACATTTAAAGAAATAGAGT 900
DB 841 TTCAAAAAAAATTAAGTTATTTTCTTATTAAGAAACATTTAAAGAAATAGAGT 900
QY 901 TGCCAGACTAGCCCTGGAATGTTTCCCAATTAATCAATCACTGATATTAATTG 960
DB 901 TGCCAGACTAGCCCTGGAATGTTTCCCAATTAATCAATCACTGATATTAATTG 960
QY 961 GCGAGCCCATTAATTAATTTAAACGAACTGAATCGAGCGAAACCAATCTGACTAT 1020
DB 961 GCGAGCCCATTAATTAATTTAAACGAACTGAATCGAGCGAAACCAATCTGACTAT 1020
QY 1021 TTCTCTAGATTAATTAAGGAGAGAGAGAGAGAAATCACTTTTAAGTCACTGTC 1080
DB 1021 TTCTCTAGATTAATTAAGGAGAGAGAGAGAGAAATCACTTTTAAGTCACTGTC 1080
QY 1081 TGAGATGTCGGGTTGGCAACGATAGCAACCTGATCATAGGCTAGGCTAGCTCA 1140
DB 1081 TGAGATGTCGGGTTGGCAACGATAGCAACCTGATCATAGGCTAGGCTAGCTCA 1140
QY 1141 GGTTCGGAGCTCTGTCGTCATCTCAATGAGCACTAGCTAGCTTTCAACCGTTC 1200
DB 1141 GGTTCGGAGCTCTGTCGTCATCTCAATGAGCACTAGCTAGCTTTCAACCGTTC 1200
QY 1200 GGTTCGGAGCTCTGTCGTCATCTCAATGAGCACTAGCTAGCTTTCAACCGTTC 1200
DB 1200 GGTTCGGAGCTCTGTCGTCATCTCAATGAGCACTAGCTAGCTTTCAACCGTTC 1200

```

```

QY 1201 TTGTTCCATGTCGCAAGCCTTGCTTATCTGAAACCAAGGATTACTACTCCCAACAT 1260
DB 1201 TTGTTCCATGTCGCAAGCCTTGCTTATCTGAAACCAAGGATTACTACTCCCAACAT 1260
QY 1261 CCATCTTACTCATGCAACTTCCATGCAACAGCAGCATATGTTCTCTGAAC 1311
DB 1261 CCATCTTACTCATGCAACTTCCATGCAACAGCAGCATATGTTCTCTGAAC 1311

RESULT 3
US-08-968-542C-1/c
; Sequence 1, Application US/08968542C
; Patent No. 5981728
; GENERAL INFORMATION:
; APPLICANT: Myers, et al.
; TITLE OF INVENTION: d(11) Codes For A No. 5981728el Starch
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McGregor & Adler, LLP
; STREET: 8011 Candle Lane
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word 6.0.1 for Macintosh
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/968,542C
; FILING DATE: No. 5981728ember 12, 1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benjamin Aaron Adler, Ph.D., J.D.
; REGISTRATION NUMBER: 35,423
; REFERENCE/DOCKET NUMBER: D6036
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713) 777-2321
; TELEFAX: (713) 777-6908
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6027 bp
; TYPE: nucleic acid
; STRANDEDNESS: double-stranded
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: cDNA to mRNA
; HYPOTHEICAL: no
; ANTI-SENSE: no
; ORIGINAL SOURCE:
; ORGANISM: maize
; TISSUE TYPE: endosperm
; IMMEDIATE SOURCE:
; LIBRARY: maize endosperm cDNA library in
; LIBRARY: gtl1
; CLONE: pmg10; pmg6aa; pmg6-2M
; US-08-968-542C-1

Query Match 5.1%; Score 66.4; DB 2; Length 6027;
Best Local Similarity 77.5%; Pred. No. 2.4e-06;
Matches 93; Conservative 0; Mismatches 26; Indels 1; Gaps 1;
QY 792 TTAAAGAGCTAGTTGGCAACCTGTTCTTCAAAAGATTTTCAAAAAA 851
DB 5606 TCTACGGGCTAGTTGGCAACCTGTTCTTCAAAAGATTTTCAAAAAA 5548
QY 852 TTAGTTATTTCTGTTATTAATAAGAAACATTAAGAAATGAGTTGCAGACTAG 911
DB 852 TTAGTTATTTCTGTTATTAATAAGAAACATTAAGAAATGAGTTGCAGACTAG 911

```

Db 5547 TTAGTTTATTTTCATTGGAAAAATTTGAATCTCTTGGAAAAATAGATTACATACATAG 5488

RESULT 4
US-09-554-467A-1/c
Sequence 1, Application US/09554467A
Patent No. 6639125
GENERAL INFORMATION:
APPLICANT: Myers, Alan M.
APPLICANT: James, Martha G.
TITLE OF INVENTION: dulli Coding for a No. 6639125el Starch Synthase and Uses
FILE REFERENCE: D6036PCT
CURRENT APPLICATION NUMBER: US/09/554,467A
CURRENT FILING DATE: 2000-05-12
PRIOR APPLICATION NUMBER: PCT/US98/24225
PRIOR FILING DATE: 1998-11-12
PRIOR APPLICATION NUMBER: US 08/062,102
PRIOR FILING DATE: 1997-11-12
NUMBER OF SEQ ID NOS: 37
SEQ ID NO 1
LENGTH: 6027
TYPE: DNA
ORGANISM: maize
FEATURE:
OTHER INFORMATION: cdna sequence corresponding to the gene encoding the
OTHER INFORMATION: starch synthase enzyme DUL.
US-09-554-467A-1

Query Match
Best Local Similarity 5.1%; Score 66.4; DB 4; Length 6027;
Matches 93; Conservative 0; Mismatches 26; Indels 1; Gaps 1;

Qy 792 TTAAAGCTAGTTTGGCAACCTGTTCTTTCCTTCAAGAAATTTGATTTTCAAAAAA 851
Db 5606 TCTACGGCTAGTTTGGAAACCCCATTT-TTCCAGAGATTTCATTTTCCAAAGAAAA 5548

Qy 852 TTAGTTTATTTCTCTTAAATAGAAAACTAGAAAAATAGTTGCGACACTAG 911
Db 5547 TTAGTTTATTTTCATTGGAAAAATTTGAATCTCTTGGAAAAATAGATTACATACATAG 5488

RESULT 5
US-08-410-784A-3/c
Sequence 3, Application US/08410784A
Patent No. 5912413
GENERAL INFORMATION:
APPLICANT: MYERS, ALAN M.
APPLICANT: JAMES, MARTHA G.
TITLE OF INVENTION: ISOLATION OF SUL, A STARCH DEBRANCHING
TITLE OF INVENTION: ENZYME, THE PRODUCT OF THE MAIZE GENE
TITLE OF INVENTION: SUGARY 1
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Weingarten, Schurgin, Gagnebin and Hayes LLP
STREET: Ten Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PasteSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/410,784A
FILING DATE: 24-MAR-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:

NAME: Heine, Ph.D., Holliday C
REGISTRATION NUMBER: 34,346
REFERENCE/DOCKET NUMBER: ISU-002XX
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-2290
TELEFAX: 617-451-0313
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2523 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
US-08-410-784A-3

Query Match
Best Local Similarity 4.5%; Score 58.6; DB 2; Length 2523;
Matches 105; Conservative 0; Mismatches 39; Indels 3; Gaps 2;

Qy 774 AAAATCTTCTGATTTTAAAGACTAGTTTGGCAACCTGTTCTTTCCTTCAAGAAATT 833
Db 279 ACATATACATATACATATATAGTCTAGTTTGCAATCCATTTATCAAAAG--TTT 222

Qy 834 TGATTTTTCAAAAAAATAGTTTATTTTCTCTTATPAATAGAAAACTAGAAAA 893
Db 221 TACATTTTTCAAAAATAGTTTATTTTCTCTTGA-AAAATGGAATTTCTCAGAAAA 163

Qy 894 ATAGAGTTGCCAGACTAGCCCTAGAT 920
Db 162 ATAGAGTTTACAACTAGCTTAAAT 136

RESULT 6
US-08-487-826B-13/c
Sequence 13, Application US/08487826B
Patent No. 5993827
GENERAL INFORMATION:
APPLICANT: Sim, Kim J.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
TITLE OF INVENTION: Wellems, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121.001CP1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550

TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 19124 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-487-826B-13

Query Match 4.3%; Score 57; DB 2; Length 19124;
Best Local Similarity 50.2%; Pred. No. 0.00078; Indels 0; Gaps 0;

Matches 141; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY 614 TAGAGTATGACGTTGAGATTTCTTTTCATTTCTGTATTTTGTATGTTTATAT 673

DB 15677 TAATGTTTTTTTTCTTCTTTGTTTTTATTTTATATCATTTTTTTTATAT 15618

QY 674 ACATTTCTCTCTTCAATAGAGTATTTCTCCGATTTTAAAAAGACTATAAGT 733

DB 15617 AAAATTTTTTTTAAATTTTTTTTGAAATCTTTTCATTTTATCTATCAAAATTTA 15558

QY 734 CATTTTATATAGACAGCATGTCGTGATTCGTCAAAATCTTCTGATTTT 793

DB 15557 TATTTATATATATTTTATTTTATTTTAAAAAATTTCTCTTTTTTTTTTTT 15498

QY 794 TAAGAGCTAGTTGGACACCTGTTCTTCAAGATTTTGAATTTTCAAAAAAAT 853

DB 15497 TTTTATTTAAATTAATTTTTTTTATATTTTCATTTTCTTTTTCATTTTAAATTA 15438

QY 854 AGTTTATTTCTCTTATTAATAATAGAAAAACACTAGAAAAA 894

DB 15437 GTTTTATATTTCTTTTAAATTAATATCATATATATATAA 15397

RESULT 7

US-09-313-294A-5397
Sequence 5397, Application US/09313294A

PATENT NO. 6476212

GENERAL INFORMATION:

APPLICANT: Laligudi, Raghunath V.

APPLICANT: Ito, Laura Y.

FILE OF INVENTION: POLYNUCLOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR

CURRENT APPLICATION NUMBER: US/09/313,294A

CURRENT FILING DATE: 1999-05-14

NUMBER OF SEQ ID NOS: 7600

SOFTWARE: PERL Program

SEQ ID NO 5397

LENGTH: 279

TYPE: DNA

ORGANISM: Zea mays

FEATURE:

NAME/KEY: misc_feature

OTHER INFORMATION: Incyte ID No. 6476212 700350078H1

NAME/KEY: unsure

LOCATION: 10, 12, 185-186, 204, 253, 274, 278

OTHER INFORMATION: a, t, c, g, or other

US-09-313-294A-5397

Query Match 4.3%; Score 55.8; DB 4; Length 279;
Best Local Similarity 69.0%; Pred. No. 0.00011; Indels 1; Gaps 1;

Matches 89; Conservative 0; Mismatches 39; Indels 1; Gaps 1;

QY 789 TTTTAAAGCTAGTTGGCAACCTGTTCTTTCAAAATTTGATTTTCAAAA 848

DB 3 TTCCTAANCCAGTTGGAAACCCCATTTTCCCAAGGTTTCAATTTTCCCAAGG 62

QY 849 AATATGTTAATTTCTCTTAT-AAATAGAAAAACCTAGAAAAATGAGTGCACA 907

DB 63 AAGTAGAACATTTTCCCTTGGAATAATGAAATCTTTGGGAAATGAGTTCCCAA 122

QY 908 CTAGCCCTA 916

DB 123 CTAGCCCTA 131

RESULT 8

US-09-902-540-1357/c

Sequence 1357, Application US/09902540

PATENT NO. 6833447

GENERAL INFORMATION:

APPLICANT: Goldman, Barry S.

APPLICANT: Hinkle, Gregory J.

APPLICANT: Slater, Steven C.

APPLICANT: Miegand, Roger C.

TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof

FILE REFERENCE: 38-10(15849)B

CURRENT APPLICATION NUMBER: US/09/902,540

CURRENT FILING DATE: 2001-07-10

PRIOR APPLICATION NUMBER: 60/217,883

NUMBER OF SEQ ID NOS: 16825

SEQ ID NO 1357

LENGTH: 612

TYPE: DNA

ORGANISM: Myxococcus xanthus

FEATURE:

NAME/KEY: unsure

LOCATION: (1) (612)

OTHER INFORMATION: unsure at all n locations

US-09-902-540-1357

Query Match 4.2%; Score 55.6; DB 4; Length 612;
Best Local Similarity 51.4%; Pred. No. 0.00047; Indels 0; Gaps 0;

Matches 127; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY 626 TTCAGATTTTCTTTTCAATCTGTATTTTGTATTTTATATATATTTCTCT 685

DB 560 TTTTAAATT 501

QY 686 CTACATATAGATTTTCTCCGATTTTAAATGACTAATGATTTTATATA 745

DB 500 TTTTATT 441

QY 746 AGAGACGATGTCGATTTCTCGTCAAAATCTTCTGATTTTAAAGCTAGT 805

DB 440 ATTTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTT 381

QY 806 TGGCAACCTGTTCTTCAAGAAATTTGATTTTCAAAAAAATGATTTATTTCT 865

DB 380 TTTTATTANTTT 321

QY 866 CTTTATA 872

DB 320 TATTTTA 314

RESULT 9

US-08-880-499-1/c

Sequence 1, Application US/08880499

PATENT NO. 6037523

GENERAL INFORMATION:

APPLICANT: Albertson, Marc C.

APPLICANT: Fox, Tim W.

APPLICANT: Carl, Garnat W.

APPLICANT: Hufman, Gary A.

TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.

```

COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/880,499
FILING DATE: CONCURRENTLY HERewith
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Sweeney, Patricia A.
REGISTRATION NUMBER: 32,733
REFERENCE/DOCKET NUMBER: 0578
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 248-4800
TELEFAX: (515) 248-4844
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1394 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-880-499-2

Query Match 4.2%; Score 55.6; DB 3; Length 1394;
Best Local Similarity 55.8%; Pred. No. 0.00064;
Matches 106; Conservative 0; Mismatches 84; Indels 0; Gaps 0

```

	Query Match	4.2%;	Score 55.6;	DB 3.	Length 1394;
	Best Local Similarity	55.8%;	Pred. No. 0.00064;		
	Matches 106;	Conservative 0;	Mismatches 84;	Indels 0;	Gaps 0
QY	761 TAGATTCGTTCCAAAAATCTTCTGAATTTTTTAAGAGCTAGTTGGCAACCCGTGTTTC				820
Dd	950 TACCACTGATTGTAATTTATTGGGAAACAATTCTAGGGCTAGTGTGCCAACTTATTTT				891
QY	821 TTTCAAAAGAAATTTGATTTTTTCAA AAAAATTAGTTATTTCTCTTATAAATAAGAA				880
Dd	890 TCATAGTGTTTCTATTTTATAAAGA AATAATACTAATTTTTTTGAAAAATCAAAA				831
QY	881 AACACTTAGAAAAATAGATTGCCAGACTAGCCCTAGCAATGTTTTCCCAATTAATTACA				940
Dd	830 TTCCTTTAAAGAAACAGGGTCCAAACTAGCTCTTAAAAAATCAGAAAGATTTTGAA				771
QY	941 TCACTGTGTA	950			
Dd	770 CGAGAACTTA	761			
RESULT 11					
US-09-806-708B-23					
; Sequence 23, Application US/09806708B					
Patent No. 6784342					

```

RESULT 11
US-09-806-708B-23
Sequence 23, Application US/09806708B
Patent No. 6784342
GENERAL INFORMATION:
APPLICANT: The University of British Columbia
TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
FILE REFERENCE: 4810-58741
CURRENT APPLICATION NUMBER: US/09/806,708B
CURRENT FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: US 60/147,133
PRIOR FILING DATE: 1999-08-04
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn version 3.0
SEQ ID NO 23
LENGTH: 1055
TYPE: DNA
ORGANISM: Artificial sequence
FEATURE:
NAME/KEY: promoter
LOCATION: (1)..(1055)
OTHER INFORMATION: consensus sequence of A.t. and L.a. FAE1 promoters
US-09-806-708B-23

Query Match          4.2%; Score 55.4; DB 4; Length 1055;
Best Local Similarity 22.5%; Pred. No. 0.00065;
Matches 182; Conservative 181; Mismatches 430; Indels 15; Gaps

```

```

QY 292 TCACCTATATATTATCTTAGATATATTATTTTGGAAAAATACAACTATAC 351
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 37 YCANNTGKRCYARWGWMTTAYMTATKMGTTGMAWTMMWMAKKTRKMCSTANNMW 96
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 352 TTTTGTAGAGGCGCTCAG-CATAGATTTCCTTAGGCGCCAGAAAAAGCAGACGAC 410
      || : : : : : : : : : : : : : : : : : : : : : : : : : :
      : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 97 TTMCTARKMTGTGRMTKTNNAATGRMTGYMTNNNGSITMWARRYTRRMWCTYAM 156
      : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 411 AAGCTAGTGTCCAGTATTTGGCACTACCCAGAAAGATTAAAAAATACCAAGTAA 470
      : : : : : : : : : : : : : : : : : : : : : : : : : :
      : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 157 MYGASWAGNASTRTTYTRWMMCKRSABARATGRARYMAYTARRGMTKMAA 216
      : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 471 CTATTCACCTCGAAAGCTATCATGTATGT---TTAAGAAACATATTAAAAACCA 527
      : : : : : : : : : : : : : : : : : : : : : : : : : :
      : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 217 YAAWTNNNNNNNAKACRKATTTGWRKSNCTCTTGTTTCAATTCMAATTCAGAWATK 276
      : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 528 TCCTCTTAAAAAACAAGCATATTTTCGAAAGAGACAATTTGTACAGTTTACAAATC 587
      : : : : : : : : : : : : : : : : : : : : : : : : : :
      : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 277 WKTWKSAAAGMTNNNNNNNNNTTTTAAATTAARMMMSATTTWMAAATSRKTWYIGR 336
      : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 588 TAAGACGCAAAATATATCGAAAGGTAAAGCTATGACGTTCAAGTTTCTTTTCATC 647
      : : : : : : : : : : : : : : : : : : : : : : : : : :
      : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 337 KTMANNNGTTCWRMAMATWMMKTKGTTWNNNGRITYGTMTKXAATTTTAAANNC 396
      : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 648 TTGTTATTTTGTATGTTTATATACATTTCTCTCTTACAAATAGAGTATTTCTT 707
      : : : : : : : : : : : : : : : : : : : : : : : : : :
      : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 397 TTAAMKTKCTMNNNTAAKATTYMATCYMKSMTNGSYRYAARYYTMAWTRRYVANN 456
      : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 708 CCGATTTTATPAAATGACTATA-----AAGTCATTTTATATATAGACGACAT 756
      : : : : : : : : : : : : : : : : : : : : : : : : : :
      : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 457 NNTKTTWAKCTWTTTKRCCCTTANNNTAAWTKSSANCTSRTRKTKMCRAGSKTASNGR 516
      : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 757 GTCTGAGATTCTGTCAAAATACTTCTGATTTTAAAGACTAGTTTGGCAACCTG 816
      : : : : : : : : : : : : : : : : : : : : : : : : : :
      : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 517 YARAYTGYKMTAMAYWCTWYVYRAGAATWAMWMTSATCYCAPAATATAGCAGAGS 576
      : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 817 TTTCTTCAAGAAATTTGATTTTTCAAAAAATAGTTTATTTCTCTTATTAAT 876
      : : : : : : : : : : : : : : : : : : : : : : : : : :
      : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 577 TAGNNNNNNNNNCAATCAATCAATCAASACAMANNCTCTYANANNATYTMNNATGCM 636
      : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 877 AGAAAACTTAGAAAAATAGAGTTGCCAGACTAGCCCTGAATGTTTCCCAATAAT 936
      : : : : : : : : : : : : : : : : : : : : : : : : : :
      : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 637 NATKTATMTNNNNNNNAGTWTNNNNNNNATASAAATTAATAATKATNTATTMAG 696
      : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 937 ACAATCACTGTATATATATTTGGCCAGCCCATTAATTTTAAACCAACGAAAT 996
      : : : : : : : : : : : : : : : : : : : : : : : : : :
      : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 697 AYABAAYTTRTANNACCTTTTNTTGTGMRNTTAARGMANNNNNNNNNNNGACMA 756
      : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 997 CGAGCGAAACCAATCGAGCTATTTCTCTAGATTAGTAAAAAGAGAGAGAGAG 1056
      : : : : : : : : : : : : : : : : : : : : : : : : : :
      : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 757 WRTTATANCNTNNNNNNNNNNNNAATTTTATTTTWRTRKANNNNNNNNNAAYGAA 816
      : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1057 AATCAGTTTAAAGTATGTCCTGAG 1084
      : : : : : : : : : : : : : : : : : : : : : : : : : :
      : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 817 KNTTWCTCKAMKAMATGAATTTNAG 844
      : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 12

```

US-09-004-056-1/c
; Sequence 1, Application US/09004056A
; Patent No. 6566586
; GENERAL INFORMATION:
; APPLICANT: Calgene LLC
; TITLE OF INVENTION: Plant Expansion Promoter Sequences
; FILE REFERENCE: 125
; CURRENT APPLICATION NUMBER: US/09/004,056A
; EARLIER FILING DATE: 1998-01-07
; EARLIER APPLICATION NUMBER: 60034914
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1

```

```

; LENGTH: 2614
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (930)
; OTHER INFORMATION: unknown nucleotide
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (947)
; OTHER INFORMATION: unknown nucleotide
; NAME/KEY: promoter
; LOCATION: (956)
; OTHER INFORMATION: unknown nucleotide
US-09-004-056-1

```

```

Query Match          4.2%; Score 55.2; DB 4; Length 2614;
Best Local Similarity 54.2%; Pred. No. 0.003;
Matches 155; Conservative 0; Mismatches 128; Indels 3; Gaps 2;

```

```

QY 611 AGTAAAGCTATGACGTTGAGATTTTCTTTTCAATCTGTTATTTGTTATGTTTAA 670
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 447 AGTTAGTTGTTATGATCAATTTTTCATATTAATTTTATTTTATTTCTTAATTTA 388
      : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 671 TATACATTTCTCTCTTCAATAGAGTATTTCTCGATTTTATATAAAGACTATA 730
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 387 TGTGCAAAATGAACCTTATTTATATATTTTAAATATATTTTGTATAAAT-TTTAA 329
      : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 731 AGTCATTTTATATATAGACGACATGTCTGATTCGTCAAAAATCTTCTGATTT 790
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 328 TATTTTCAATATATATTTTCAGAAACAAATATTTTGAATTAACGAATTTTGAGATT 269
      : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 791 TTTAAGACTAGTTTGGCAACCTGTTCTTCAAAAGATTTTGAATTTTCAAAAAA 850
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 268 TTTAAAGTTAGTATTTTAAATTTCTTATTTATATATGAATAT-TTATTTAGAAAAA 211
      : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 851 ATTAGTTATTTCTCTTATATAAATAGAAAAACCTTAGAAAAAT 896
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 210 ATTAGTTAATGAATTTAAATTTTAAATAATTTAAAAAATA 165
      : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 13

```

US-09-640-173-53
; Sequence 53, Application US/09640173
; Patent No. 6613515
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: OVARIAN TUMOR SEQUENCES AND
; FILE REFERENCE: 210121.484C2
; CURRENT APPLICATION NUMBER: US/09/640,173
; EARLIER FILING DATE: 2000-08-15
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 53
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1) ... (396)
; OTHER INFORMATION: n = A,T,C or G
US-09-640-173-53

```

```

Query Match          4.2%; Score 55; DB 4; Length 396;
Best Local Similarity 43.8%; Pred. No. 0.00056;
Matches 139; Conservative 0; Mismatches 178; Indels 0; Gaps 0;

```

```

QY 632 TTTTCTTTTCATCTGTATTTTGTATGTTTATATACATTTTCTCTCTAC 691
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 11 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 70
      : : : : : : : : : : : : : : : : : : : : : : : : : :

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: September 15, 2005, 07:10:11; Search time 853.804 Seconds
(without alignments)
10230.248 Million cell updates/sec

Title: US-10-713-381-2_COPY_1_1311

Perfect score: 1311
Sequence: 1 cccatgctgctctatgaataa.....cgcacatagctctcgaac 1311

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 7389322 seqs, 33128559 residues

Total number of hits satisfying chosen parameters: 14778644

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

Published Applications NA:
1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/prodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/prodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/prodata/2/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/prodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/prodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/prodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/prodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/prodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/prodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/prodata/2/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/prodata/2/pubpna/US10G_PUBCOMB.seq:*
20: /cgn2_6/prodata/2/pubpna/US10H_PUBCOMB.seq:*
21: /cgn2_6/prodata/2/pubpna/US10I_PUBCOMB.seq:*
22: /cgn2_6/prodata/2/pubpna/US10J_NEW_PUB.seq:*
23: /cgn2_6/prodata/2/pubpna/US11A_PUBCOMB.seq:*
24: /cgn2_6/prodata/2/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1311	100.0	1394	20	US-10-713-381-1
2	1311	100.0	1394	20	US-10-713-381-2
3	160.6	12.3	255	20	US-10-713-381-9
4	146	11.1	158	20	US-10-713-381-3
5	80.8	6.2	1261	20	US-10-425-115-134230
6	74	5.6	320	18	US-10-425-114-23340
7	74	5.6	624	18	US-10-425-114-16264

8	73	5.6	1326	20	US-10-425-115-141826	Sequence 141826, A
9	72.8	5.6	2445	18	US-10-425-114-32493	Sequence 32493, A
10	72.8	5.6	2729	20	US-10-425-115-83293	Sequence 83293, A
11	72.4	5.5	1203	20	US-10-425-115-51470	Sequence 51470, A
12	71.8	5.5	928	18	US-10-425-114-17816	Sequence 17816, A
13	71.8	5.5	1123	20	US-10-425-115-142853	Sequence 142853, A
14	71.8	5.5	2537	18	US-10-425-114-31957	Sequence 31957, A
15	71.8	5.5	2537	20	US-10-425-115-52216	Sequence 52216, A
16	71.8	5.5	3607	18	US-10-425-114-31061	Sequence 31061, A
17	71.8	5.5	3691	20	US-10-425-115-52219	Sequence 52219, A
18	70.2	5.4	960	14	US-10-198-846-6381	Sequence 6381, Ap
19	69.8	5.3	13400	21	US-10-343-477A-63	Sequence 63, Ap1
20	69.6	5.3	724	20	US-10-425-115-87756	Sequence 87756, A
21	69.2	5.3	527	18	US-10-425-114-4041	Sequence 4041, Ap
22	68.8	5.2	610	20	US-10-425-115-47165	Sequence 47165, A
23	68.4	5.2	1215	18	US-10-425-114-24656	Sequence 24656, A
24	68.4	5.2	2863	20	US-10-425-115-75310	Sequence 75310, A
25	67.8	5.2	512	20	US-10-425-115-49781	Sequence 49781, A
26	67.8	5.2	928	18	US-10-425-114-17816	Sequence 17816, A
27	67.8	5.2	1123	20	US-10-425-115-142853	Sequence 142853, A
28	67.4	5.1	1260	18	US-10-425-114-30881	Sequence 30881, A
29	67.4	5.1	1411	20	US-10-425-115-149304	Sequence 149304, A
30	67.2	5.1	2232	18	US-10-425-114-16607	Sequence 16607, A
31	67.2	5.1	2249	18	US-10-425-114-20264	Sequence 20264, A
32	66.4	5.1	6027	18	US-10-634-262-1	Sequence 1, Appli
33	66.4	5.1	6027	19	US-10-109-048-1145	Sequence 1145, Ap
34	66.4	5.1	6051	20	US-10-425-115-41589	Sequence 41589, A
35	66.2	5.0	439	9	US-09-960-352-1009	Sequence 1009, Ap
36	66	5.0	1376	18	US-10-425-115-177935	Sequence 177935, Ap
37	65.6	5.0	721	20	US-10-425-114-777	Sequence 777, App
38	65.6	5.0	2274	18	US-10-425-114-830	Sequence 830, App
39	65.6	5.0	2632	20	US-10-425-114-162957	Sequence 162957, A
40	65.6	5.0	8056	20	US-10-473-126-386	Sequence 386, App
41	65.6	5.0	1364	20	US-10-425-115-150820	Sequence 150820, A
42	65	5.0	833	20	US-10-425-115-30824	Sequence 30824, A
43	64.8	4.9	1838	20	US-10-425-115-21012	Sequence 21012, A
44	64.8	4.9	581	20	US-10-425-115-46324	Sequence 46324, A
45	64.2	4.9				

ALIGNMENTS

RESULT 1
US-10-713-381-1
; Sequence 1, Application US/10713381
; Publication No. US2004022131A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-1

Query Match 100.0%; Score 1311; DB 20; Length 1394;
Best Local Similarity 100.0%; Pred. No. 5e-275;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCATGCTGCTCTATGAATAAGATGATGATGCTCTATATCCGTTTCTTAGGCTCC 60

```
Db 1 CCAATGGTGTCTATGAAAAAGATGATCAATGTGTCTATATCCGTTTTCTTAAGGTC 60
QY 61 CTTCTTCTGCTTATTAAGTACTGAAATCGGGGTTACAAAACTTCACAGGGTGCATGAT 120
Db 61 CTTCTTCTGCTTATTAAGTACTGAAATCGGGGTTACAAAACTTCACAGGGTGCATGAT 120
QY 121 CTCATGTTTCCACTTCCCACTCGGGTGGACATTTCTTGGATGTGGTGGTCCCAT 180
Db 121 CTCATGTTTCCACTTCCCACTCGGGTGGACATTTCTTGGATGTGGTGGTCCCAT 180
QY 181 CTGACCGAGGCCCATAGACACCTTTGGGACACCCATCAAGGGCTTTGGATGGCCCA 240
Db 181 CTGACCGAGGCCCATAGACACCTTTGGGACACCCATCAAGGGCTTTGGATGGCCCA 240
QY 241 CGAGACGTATCGGGTGTGTGATCCAGGGGATATATGTCCCAATGTCTACTATTA 300
Db 241 CGAGACGTATCGGGTGTGTGATCCAGGGGATATATGTCCCAATGTCTACTATTA 300
QY 301 TTATTAATCTTGTAGATATTAATTTTGGAAAAATACAACTTAATACCTTTGTGTA 360
Db 301 TTATTAATCTTGTAGATATTAATTTTGGAAAAATACAACTTAATACCTTTGTGTA 360
QY 361 GGGCTCAGCATAGATTTTGCCTTAAGGCCAGAAAATGCGAGGACCGCATGTCTAGTG 420
Db 361 GGGCTCAGCATAGATTTTGCCTTAAGGCCAGAAAATGCGAGGACCGCATGTCTAGTG 420
QY 421 TCCACTATTTGGCATACCCAGAACAAAGTTTAAAAAATACCAAGTAATCACTACT 480
Db 421 TCCACTATTTGGCATACCCAGAACAAAGTTTAAAAAATACCAAGTAATCACTACT 480
QY 481 CGAAGCTATCATGTATGTTTAAAGAAACATCTATTTAAACCAAGATCCTCTTAAAAA 540
Db 481 CGAAGCTATCATGTATGTTTAAAGAAACATCTATTTAAACCAAGATCCTCTTAAAAA 540
QY 541 CAAACATATTTGAAAAGACAAATTAATGTACAGTTTACAAACATCTAAGACGACAA 600
Db 541 CAAACATATTTGAAAAGACAAATTAATGTACAGTTTACAAACATCTAAGACGACAA 600
QY 601 TTATATCGAAAGGTAGCTATGACGTTCAAGATTTTCTTTCAATCTTGTATTTGTT 660
Db 601 TTATATCGAAAGGTAGCTATGACGTTCAAGATTTTCTTTCAATCTTGTATTTGTT 660
QY 661 ATGTGTTTATATACATTTCTCTCTTACAAATAGATGATTTCTTCGATTTTAA 720
Db 661 ATGTGTTTATATACATTTCTCTCTTACAAATAGATGATTTCTTCGATTTTAA 720
QY 721 ATGACTATAAAGTCATTTTATATAAAGACGCAATGTGTAGATTTCTGTTCAAAAATC 780
Db 721 ATGACTATAAAGTCATTTTATATAAAGACGCAATGTGTAGATTTCTGTTCAAAAATC 780
QY 781 TTTCTGATTTTAAAGAGCTATTTGGCAACCTGTCTTCTTCAAGAAATTTGATTT 840
Db 781 TTTCTGATTTTAAAGAGCTATTTGGCAACCTGTCTTCTTCAAGAAATTTGATTT 840
QY 841 TTCAAAAAAATAGTTTATTTCTCTTATAAATAGAAAAACACTTAGAAAAATAGAT 900
Db 841 TTCAAAAAAATAGTTTATTTCTCTTATAAATAGAAAAACACTTAGAAAAATAGAT 900
QY 901 TGGCAGACTAGCCCTAGAAATGTTTCCCAATAAATTAACAATCACTGTGTATTAATTTG 960
Db 901 TGGCAGACTAGCCCTAGAAATGTTTCCCAATAAATTAACAATCACTGTGTATTAATTTG 960
QY 961 GCGACGCCCATTAATTTTAAACCGAAATCGAAGCGAAACCAATCTGAGCTAT 1020
Db 961 GCGACGCCCATTAATTTTAAACCGAAATCGAAGCGAAACCAATCTGAGCTAT 1020
QY 1021 TTCTCTAGATTAGTAAAGAGAGAGAGAGAGAAATCAGTTTAAAGTCAATGTGCC 1080
Db 1021 TTCTCTAGATTAGTAAAGAGAGAGAGAGAGAAATCAGTTTAAAGTCAATGTGCC 1080
QY 1081 TGAAGTGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGTCAATAGTGTGCTACGTA 1140
Db 1081 TGAAGTGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGTCAATAGTGTGCTACGTA 1140
```

```
Db 1081 TGAAGTGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGTCAATAGTGTGCTACGTA 1140
QY 1141 GGTGGGAGCTCTGTGTGATCTTCAATGGATATCTATACCTGTTTCAACCGTTGTC 1200
Db 1141 GGTGGGAGCTCTGTGTGATCTTCAATGGATATCTATACCTGTTTCAACCGTTGTC 1200
QY 1201 TTGTTCAATGCTCAAGCCTTGCTAATTTCTGAACCAAGAGATACCTACTCCAAACAT 1260
Db 1201 TTGTTCAATGCTCAAGCCTTGCTAATTTCTGAACCAAGAGATACCTACTCCAAACAT 1260
QY 1261 CCATCTTAATCAATGCACTTCCATGCAACGACGACATATGTTTCTGAAAC 1311
Db 1261 CCATCTTAATCAATGCACTTCCATGCAACGACGACATATGTTTCTGAAAC 1311

RESULT 2
US-10-713-381-2
; Sequence 2, Application US/10713381
; Publication No. US2004022131A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAAT, CARL W.
; APPLICANT: HUFEMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE--PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: USING SAME
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/860,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patencin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-2

Query Match 100.0%; Score 1311; DB 20; Length 1394;
Best Local Similarity 100.0%; Pred. No. 5e-275;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAATGGTGTCTATGAAAAAGATGATCAATGTGTCTATATCCGTTTTCTTAAGGTC 60
Db 1 CCAATGGTGTCTATGAAAAAGATGATCAATGTGTCTATATCCGTTTTCTTAAGGTC 60
QY 61 CTTCTTCTGCTTATTAAGTACTGAAATCGGGGTTACAAAACTTCACAGGGTGCATGAT 120
Db 61 CTTCTTCTGCTTATTAAGTACTGAAATCGGGGTTACAAAACTTCACAGGGTGCATGAT 120
QY 121 CTCATGTTTCCACTTCCCACTCGGGTGGACATTTCTTGGATGTGGTGGTCCCAT 180
Db 121 CTCATGTTTCCACTTCCCACTCGGGTGGACATTTCTTGGATGTGGTGGTCCCAT 180
QY 181 CTGACCGAGGCCCATAGACACCTTTGGGACACCCATCAAGGGCTTTGGATGGCCCA 240
Db 181 CTGACCGAGGCCCATAGACACCTTTGGGACACCCATCAAGGGCTTTGGATGGCCCA 240
QY 241 CGAGACGTATCGGGTGTGTGATCCAGGGGATATATGTCCCAATGTCTACTATTA 300
Db 241 CGAGACGTATCGGGTGTGTGATCCAGGGGATATATGTCCCAATGTCTACTATTA 300
QY 301 TTATTAATCTTGTAGATATTAATTTTGGAAAAATACAACTTAATACCTTTGTGTA 360
Db 301 TTATTAATCTTGTAGATATTAATTTTGGAAAAATACAACTTAATACCTTTGTGTA 360
QY 361 GGGCTCAGCATAGATTTTGCCTTAAGGCCAGAAAATGCGAGGACCGCATGTCTAGTG 420
Db 361 GGGCTCAGCATAGATTTTGCCTTAAGGCCAGAAAATGCGAGGACCGCATGTCTAGTG 420
QY 421 TCCACTATTTGGCATACCCAGAACAAAGTTTAAAAAATACCAAGTAATCACTACT 480
Db 421 TCCACTATTTGGCATACCCAGAACAAAGTTTAAAAAATACCAAGTAATCACTACT 480
```

```
Db 421 TCCACTATTGGCACTACCCAGAACAGATTTTAAAAAATTAACCAATTAATCACT 480
Qy 481 CGAAGCTATCATGTAAATGTTTAAAGAAACATCTATTAAACCAAGATCTCTTAAAAA 540
Db 481 CGAAGCTATCATGTAAATGTTTAAAGAAACATCTATTAAACCAAGATCTCTTAAAAA 540
Qy 541 CAAGCATATTTCGAAAGAGACAAATTATGTTACGTTTACAAACATTAAGAGACAA 600
Db 541 CAAGCATATTTCGAAAGAGACAAATTATGTTACGTTTACAAACATTAAGAGACAA 600
Qy 601 TTAATATGAAAGGTAAAGCTATGACGTTCAAGATTTTCTTTTCACTCTGTTATTTGTT 660
Db 601 TTAATATGAAAGGTAAAGCTATGACGTTCAAGATTTTCTTTTCACTCTGTTATTTGTT 660
Qy 661 ATTGTTTTTATATCATTTTCTCTCTTACAAATAGATGATTTTCTCCGATTTTATAA 720
Db 661 ATTGTTTTTATATCATTTTCTCTCTTACAAATAGATGATTTTCTCCGATTTTATAA 720
Qy 721 ATGACTATAAGTCATTTTATATATTAAGAGACGATGCTGATGTTCTGTTCAAAAATC 780
Db 721 ATGACTATAAGTCATTTTATATTAAGAGACGATGCTGATGTTCTGTTCAAAAATC 780
Qy 781 TTTCTGATTTTATTAAGAGCTAGTTTGGCAACCTGTTCTTTCAAGAAATTTGATTTT 840
Db 781 TTTCTGATTTTATTAAGAGCTAGTTTGGCAACCTGTTCTTTCAAGAAATTTGATTTT 840
Qy 841 TTCAAAAAAATAGTTATTTTCTCTTATTAATAAGAAACATTAAGAAATATAGAT 900
Db 841 TTCAAAAAAATAGTTATTTTCTCTTATTAATAAGAAACATTAAGAAATATAGAT 900
Qy 901 TGGCAAGCTAGCCCTAGAAATGTTTCCCAATTAATTAATCACTGTTGTTATTTGTT 960
Db 901 TGGCAAGCTAGCCCTAGAAATGTTTCCCAATTAATTAATCACTGTTGTTATTTGTT 960
Qy 961 GCCAGCCCCATTAATTTTAAACCGAACTGAATTCGACGAAACCAATCTGAGCTAT 1020
Db 961 GCCAGCCCCATTAATTTTAAACCGAACTGAATTCGACGAAACCAATCTGAGCTAT 1020
Qy 1021 TTCTCTAGATTAGTAAAGAGAGAGAGAGAGAGAGAAATCAGTTTAACTCATTTGCC 1080
Db 1021 TTCTCTAGATTAGTAAAGAGAGAGAGAGAGAGAGAAATCAGTTTAACTCATTTGCC 1080
Qy 1081 TGAGATGTGGGTTTGGCAACGATAGCCAGCTAATATAGCTCATAGGCTTAGCTCA 1140
Db 1081 TGAGATGTGGGTTTGGCAACGATAGCCAGCTAATATAGCTCATAGGCTTAGCTCA 1140
Qy 1141 GGTTCGGAGCTCTCGTGTATCTCAATGAGCATGCTACTACTGCTTGTCAACGTTGCTC 1200
Db 1141 GGTTCGGAGCTCTCGTGTATCTCAATGAGCATGCTACTACTGCTTGTCAACGTTGCTC 1200
Qy 1201 TTGTTTCATGCTCGAAGCCCTGCTATTTCTGAAACCAAGAGATACCTATCCCAACAT 1260
Db 1201 TTGTTTCATGCTCGAAGCCCTGCTATTTCTGAAACCAAGAGATACCTATCCCAACAT 1260
Qy 1261 CCATCTTACATGCAACTTCCATGCAACACGCACTAATGTTTCTGAAAC 1311
Db 1261 CCATCTTACATGCAACTTCCATGCAACACGCACTAATGTTTCTGAAAC 1311

RESULT 3
US-10-713-381-9
; Sequence 9, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFEMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
```

```
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 255
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-9
Query Match 12.3%; Score 160.6; DB 20; Length 255;
Best Local Similarity 97.6%; Pred. No. 6.1e-25;
Matches 163; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Db 1145 CGGCACTCTCGTGTATCTCAATGCAATGCTATGCTTGTCAACGTTGCTGTT 1204
Qy 5 CCGCGATCCCGTGTATCTCAATGCAATGCTATGCTTGTCAACGTTGCTGTT 64
Db 1205 TCCATGCTCCAGCTTGTCTTATTTGAAACCAAGAGATACCTATCCCAACATCAT 1264
Qy 65 TCCATGCTCCAGCTTGTCTTATTTGAAACCAAGAGATACCTATCCCAACATCAT 124
Db 1265 CTACTCATGCAACTTCCATGCAACACGCAATATGTTTCTGAAAC 1311
Qy 125 CTACTCATGCAACTTCCATGCAACACGCAATATGTTTCTGAAAC 171

RESULT 4
US-10-713-381-3
; Sequence 3, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFEMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 158
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-3
Query Match 11.1%; Score 146; DB 20; Length 158;
Best Local Similarity 99.4%; Pred. No. 7.3e-22;
Matches 157; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 1155 CGTGTATCTCAATGCAATGCTATGCTTGTCAACGTTGCTGTTTCTGATCTGTC 1213
Db 1 CGTGTATCTCAATGCAATGCTATGCTTGTCAACGTTGCTGTTTCTGATCTGTC 60
Qy 1214 CAAGCTTGTCTATTTGAAACCAAGAGATACCTATCTCCCAACATCATCTTACTCAT 1273
Db 61 CAAGCTTGTCTATTTGAAACCAAGAGATACCTATCTCCCAACATCATCTTACTCAT 120
Qy 1274 GCAACTTCATGCAACACGCACTATGTTTCTGAAAC 1311
Db 121 GCAACTTCATGCAACACGCACTATGTTTCTGAAAC 158

RESULT 5
US-10-425-115-134230
```



```
/ Sequence 134230, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(53222)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 134230
/ LENGTH: 1261
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MRT4577_53903C.1
US-10-425-115-134230
```

```
Query Match          6.2%; Score 80.8; DB 20; Length 1261;
Best Local Similarity 78.2%; Pred. No. 3.1e-07;
Matches 97; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
```

```
Qy 792 TTTAAGAGCTAGTTGGCAACCTGTTCTTCAAGAAATTTGATTTTCAAAAAA 851
    |||||
Db 999 TCTTAAGGCTAGTTGGAACTTATTTTAAAGATTTCTTTTAAAGATTA 1058
    |||||

Qy 852 TTAGTTATTTCTCTTATTAATAATGAAAAACATTGAAATAGAGTCCAGACTAG 911
    |||||
Db 1059 TTAGTTATTTCTCTTGAAGAAATGAAATCTCTTGAAGAAATGAGTTGCTAAACTAG 1118
    |||||

Qy 912 CCCT 915
    |||||
Db 1119 CCCT 1122
    |||||
```

```
RESULT 6
US-10-425-114-23340
/ Sequence 23340, Application US/10425114
/ Publication No. US20040034888A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jingdong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E.
/ APPLICANT: Tabaska, Jack E.
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(5313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 23340
/ LENGTH: 320
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: LIB3595-039-G8_FLI
US-10-425-114-23340
```

```
Query Match          5.6%; Score 74; DB 18; Length 320;
Best Local Similarity 73.1%; Pred. No. 4.8e-06;
Matches 95; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
```

```
Qy 791 TTTAAGAGCTAGTTGGCAACCTGTTCTTCAAGAAATTTGATTTTCAAAAAA 850
    |||||
Db 62 TTTTGGGCTAGTTGGAATCTCATTTTTCAGAGATTTTATTTCCAAAGGA 121
    |||||

Qy 851 ATTAGTTATTTCTCTTATTAATAATGAAAAACATTGAAATAGAGTCCAGACTA 910
    |||||
```

```
Db 122 ATTAGTTATTTCTCTTGAAGAAATGAAATTCCTTGGAAATTAGAGTTCTAAACGA 181
Qy 911 GCCCTAGAT 920
    |||||
Db 182 GCCCTAATT 191
    |||||
```

```
RESULT 7
US-10-425-114-16264
/ Sequence 16264, Application US/10425114
/ Publication No. US20040034888A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jingdong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E.
/ APPLICANT: Tabaska, Jack E.
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(5313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 16264
/ LENGTH: 624
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: LIB3062-023-G10_FLI
US-10-425-114-16264
```

```
Query Match          5.6%; Score 74; DB 18; Length 624;
Best Local Similarity 73.1%; Pred. No. 6.7e-06;
Matches 95; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
```

```
Qy 791 TTTAAGAGCTAGTTGGCAACCTGTTCTTCAAGAAATTTGATTTTCAAAAAA 850
    |||||
Db 374 TTTTGGGCTAGTTGGAATCTCATTTTTCAGAGATTTTATTTCTTAAGGA 433
    |||||

Qy 851 ATTAGTTATTTCTCTTATTAATAATGAAAAACATTGAAATAGAGTCCAGACTA 910
    |||||
Db 434 ATTAGTTATTTCTCTTGAAGAAATGAAATCCCTTGGAAATAGAGTTCTTAACGA 493
    |||||

Qy 911 GCCCTAGAT 920
    |||||
Db 494 GCCCTAATT 503
    |||||
```

```
RESULT 8
US-10-425-115-141826
/ Sequence 141826, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(53222)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 141826
/ LENGTH: 1326
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MRT4577_60829C.1
US-10-425-115-141826
```

```
Query Match          5.6%; Score 73; DB 20; Length 1326;
```


Best Local Similarity 69.1%; Pred. No. 1.6e-05;
Matches 114; Conservative 0; Mismatches 50; Indels 1; Gaps 1;

QY 756 TGTCTAGATTCCTGTCGTAAGAAATCTTGTGATTTTAAAGACCTAGTTGGCAACCT 815

DB 1090 TGAAGTGGTGGTGTCTGCTGATTTTCACTTTAACTGTAAGACTAATTTGGGAACCA 1149

QY 816 GTTCTTCAAGAAATTTTGTATTTTCAAAAAAATTAAGTTATTTTCTTTATATAA 875

DB 1150 ATTT-TCACACTGATTTTCAATTTTCTTAAGAAATAAGTTCAATTTCCCTTGAGAAA 1208

QY 876 TAGAAACACTAGAAAAATAGAGTTGCGCACTAGCCCTAGAT 920

DB 1209 TAGAATCCCTTAGAAAAAATAAGTTCCAACTAGCCCTTAAT 1253

RESULT 9
US-10-425-114-32493

Publication NO. US20040034888A1

GENERAL INFORMATION:

APPLICANT: Liu, Jindong

APPLICANT: Zhou, Yihua

APPLICANT: Kovalic, David K.

APPLICANT: Screen, Steven E.

APPLICANT: Tabaka, Jack E.

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

FILE REFERENCE: 38-21(5313)B

CURRENT APPLICATION NUMBER: US/10/425, 114

NUMBER OF SEQ ID NOS: 2003-04-28

LENGTH: 2445

TYPE: DNA

ORGANISM: Zea mays

FEATURE:

OTHER INFORMATION: Clone ID: UC-ZMFLB73402B09_FLI

US-10-425-114-32493

Query Match 5.6%; Score 72.8; DB 18; Length 2445;

Best Local Similarity 73.6%; Pred. No. 2.3e-05;

Matches 106; Conservative 0; Mismatches 37; Indels 1; Gaps 1;

QY 792 TTTAAGCTAGTTGGCAACCCCTGTTCTT-TCAAGAATTTTGAATTTTCAAAAAA 850

DB 1623 TCTTAGGCTAGTTGGAACCTTTTTCCTCAAAAGATTTTCAATTTTCAAGAAA 1682

QY 851 ATTAGTTATTTCTCTTTATTAATAATGAAAAACATTAGAAAAATAGAGTTGCCAGCTA 910

DB 1683 ATTAGTTCATTTTCTCTGAGAAAAATAGAAATCCCTTAATAAATAGTGTTCATAA 1742

QY 911 GCCCTAGATGTTTCCCATATA 934

DB 1743 GCCCTTAATGTTTTCATGAA 1766

SEQ ID NO 83293

LENGTH: 2729

TYPE: DNA

ORGANISM: Zea mays

FEATURE:

NAME/KEY: unsure

LOCATION: (1)..(2729)

OTHER INFORMATION: unsure at all n locations

OTHER INFORMATION: Clone ID: MRT4577_175978C.1

US-10-425-115-83293

Query Match 5.6%; Score 72.8; DB 20; Length 2729;

Best Local Similarity 60.7%; Pred. No. 2.5e-05;

Matches 136; Conservative 0; Mismatches 87; Indels 1; Gaps 1;

QY 792 TTTAAGCTAGTTGGCAACCCCTGTTCTTCAAGAATTTTGAATTTTCAAAAAA 851

DB 2523 TTTAAGGCTAGTTAGAA-CTCATTTTCTTAAGATTTTCTATTTTCCAGAAAA 2465

QY 852 TTAGTTATTTCTCTTTAATAAATGAAAAACATTAGAAAAATAGAGTTGCCAGCTAG 911

DB 2464 ATATTTCATTTCCCTTGAATAATAGAAATACCTTGAAAAATAGAGTTCCAACTAG 2405

QY 912 CCTAGATGTTTCCCATTAATTCATCACTGTGTATTAATTTGGCCAGCCCAT 971

DB 2404 CTTCAAAATTTCTCTCAAGTATTTCTTATAGAGATATGTATACATGCGAGCT 2345

QY 972 AATTATTAACCGAACTGAATGAGCGAAACCAATCTGA 1015

DB 2344 TGAGTATCATGACCTAAATGTAAGAGGTAAATCAATGTTA 2301

RESULT 11
US-10-425-115-51470/c

Publication NO. US20040214272A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

FILE REFERENCE: 38-21(5322)B

CURRENT APPLICATION NUMBER: US/10/425, 115

NUMBER OF SEQ ID NOS: 2003-04-28

LENGTH: 1203

TYPE: DNA

ORGANISM: Zea mays

FEATURE:

NAME/KEY: unsure

LOCATION: (1)..(1203)

OTHER INFORMATION: unsure at all n locations

OTHER INFORMATION: Clone ID: MRT4577_146936C.1

US-10-425-115-51470

Query Match 5.5%; Score 72.4; DB 20; Length 1203;

Best Local Similarity 76.1%; Pred. No. 2e-05;

Matches 102; Conservative 0; Mismatches 31; Indels 1; Gaps 1;

QY 787 ATTTTAAAGCTAGTTGGCAACCCCTGTTCTTCAAGAATTTTGAATTTTCAAA 846

DB 1189 ACTTATTAAGGCTAGTGGGGA-CATATTTTCCAAAAAGATTTCTATTTCTTAA 1131

QY 847 AAAATTAAGTTATTTCTCTTATTAATAATGAAAAACATTAGAAAAATAGAGTTGCCAG 906

DB 1130 GAAATTAAGTTATTTCTCTTGAATAAATAATCCGTAGAAAAATAGAGTTCCAA 1071

QY 907 ACTAGCCCTAGAT 920


```

; ORGANISM: zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_147620C.1
US-10-425-115-52216

```

```

Query Match          5.5%; Score 71.8; DB 20; Length 2537;
Best Local Similarity 73.4%; Pred. No. 3.9e-05;
Matches 105; Conservative 0; Mismatches 37; Indels 1; Gaps 1;

```

```

QY      792 TTTAAGAGCTAGTTGGCAACCTGTTCTTCAAGAAATTTGATTTTTCAAAAAAA 851
        |||||
DB      1799 TCTTAGGGCTAGTTGGAACCT-TTTTCCCAAAAGATTTCATTTTCAAAGAAAA 1857
        |||||
QY      852 TTAGTTATTTCTCTTTATAAATAAGAAAACTTAGAAAAATAGAGTCCAGACTAG 911
        |||||
DB      1858 TTAGTTCATTTTCTTGAGAAAAATAGGAATCCCTTAAAAAATAGTGTTCAAACTAG 1917
        |||||
QY      912 CCTAGATGTTTCCCAATPAA 934
        |||||
DB      1918 CCTTAATGTTTTCATGAA 1940
        |||||

```

```

Search completed: September 15, 2005, 20:45:54
Job time : 854.804 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using bw model

Run on: September 15, 2005, 16:07:32 ; Search time 4088.83 Seconds
(Without alignments)
12204.542 Million cell updates/sec

Title: US-10-713-381-2_COPY_1_1311

Perfect score: 1311
Sequence: 1 ccacgtgcctctacgaagaa.....cgacacatgcttcctgaac 1311

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_nuc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_g881:*
9: gb_g882:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	908	69.3	963	9	CC656933 OGMDO20TM
2	679	51.8	915	9	CG224225 OG1AG08TV
3	419.8	32.0	687	9	CC656933 OGMDO20TV
4	96	7.3	715	9	CG252571 CG4BB05TC
5	92	7.0	967	9	CL235046 ZMMBB057
6	88	6.7	814	9	CG048704 PU1C019TB
7	87.8	6.7	754	9	CG414922 ZMMBB029
8	86.8	6.6	950	8	CC439901 PUHRA15TB
9	86.6	6.6	652	8	CC384247 PUHOC67TB
10	86.6	6.6	797	8	CC400575 PUHLU61TD
11	86.6	6.6	820	8	CC400574 PUHLU61TB
12	85.4	6.5	471	9	CG103452 PUJBR19TB
13	85.2	6.5	765	9	CG082135 PUFOX12TD
14	85.2	6.5	781	9	CG630219 OGUCC53TV
15	85.2	6.5	815	9	CG349565 OG0F183TH
16	85.2	6.5	834	9	CC630210 OGUCC53TH
17	85	6.5	789	8	CC433618 PUHHP17TD
18	84.6	6.5	1092	8	AL175696 Tetraodon
19	84	6.4	793	8	BZ816381 PUFBAB6TD
20	84	6.4	1078	9	CL957678 ZMMBB001
21	83.8	6.4	530	9	CG201774 PUICH24TB
22	83.6	6.4	722	9	CG333914 OG0AD14TH
23	83.6	6.4	722	9	CG333929 OG0AD14TV
24	83.6	6.4	781	9	CG034985 PUIGR68TB

25	83.6	6.4	861	8	BZ797976
26	83.6	6.4	865	8	CC430754 PUHEB05TB
27	81.8	6.2	861	9	CG102092 PUFTW94TD
28	81.8	6.2	947	8	CC435780 PUHNS02TD
29	81.8	6.2	981	8	BZ784278 PUFHV09TD
30	81.8	6.2	1017	9	CC620594 OGUCC62TV
31	81.6	6.2	649	9	CC613918 OGJAG88TV
32	81.6	6.2	733	8	BZ778636 1h02f10.g
33	81.6	6.2	925	9	CG071791 PUBK02TB
34	81.6	6.2	960	8	BZ676889 PUBIG17TD
35	81.6	6.2	994	8	CC003943 CUDJN65TD
36	81.6	6.2	1016	9	CL996481 ZMMBBF000
37	81.6	6.2	1022	9	CL984151 ZMMBBH000
38	81.4	6.2	999	8	CC385762 PUHFK14TD
39	81.2	6.2	756	9	CG221693 OGMWC04TH
40	81	6.2	1101	9	AL078714 Drosophila
41	80.6	6.1	699	8	BZ996930 PUJGIR50TB
42	80.6	6.1	832	9	CC676387 OGKAS81TH
43	80.6	6.1	839	9	CG246849 OGMA122TH
44	80.6	6.1	843	9	CG254365 OGMFV41TH
45	80.6	6.1	852	8	CC385881 PUHMB21TD

ALIGNMENTS

RESULT 1	CC656933	963 bp	DNA	linear	GSS 19-JUN-2003
LOCUS	OGMDO20TM ZM 0.7 1.5 KB	Zea mays	genomic clone	ZMMBMA0554D15,	
DEFINITION	genomic survey sequence.				
ACCESSION	CC656933	GI:32060225			
VERSION	CC656933.1				
KEYWORDS	GSS.				
SOURCE	Zea mays				
ORGANISM	Zea mays				
REFERENCE	Whitelaw,C.A., Quackenbush,J., Van Aken,S., Uteerback,T., Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea. 1 (bases 1 to 963)				
AUTHORS	Whitelaw,C.A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.				
TITLE	Consortium for Maize Genomics				
JOURNAL	Unpublished (2002)				
COMMENT	Other GSSes: OGMDO20TV Contact: Cathy Whitelaw TIGR				
FEATURES	source				
source	Location/Qualifiers				
	1..963				
	/organism="Zea mays"				
	/mol_type="genomic DNA"				
	/strain="B73"				
	/db_xref="taxon:4577"				
	/clone_lib="ZMMBMA0554D15"				
	/clone_lib="ZM_0.7_1.5_KB"				
	/note="vector: pBCSK-7 Site 1: HincII, 0.7-1.5 kb methylation filtered genomic DNA library"				
ORIGIN					
Query Match	69.3% Score 908; DB 9; Length 963;				
Best Local Similarity	100.0%; Pred. No. 7.8e-180; Indels 0; Gaps 0;				
Matches 908; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
404 ACCAGCCATGTCTAGTCTCACTATTTGGACATCCAGACAAATTTAAAAATAACC 463					
DB 1 ACCAGCCATGTCTAGTCTCACTATTTGGACATCCAGACAAATTTAAAAATAACC 60					

RESULT 2	
CG224225	
LOCUS	CG224225
DEFINITION	CG224225 915 bp DNA linear GSS 22-AUG-2001
ACCESSION	CG224225
VERSION	CG224225.1
KEYWORDS	Genomic survey sequence.
SOURCE	CG224225.1 GI:34124113
ORGANISM	GSS.
	Zea mays
	Zea mays
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

REFERENCE	TITLE	JOURNAL	COMMENT
Spentatophyes; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD			
clade: Panicoidae; Andropogoneae; Zea.			
1 (bases 1 to 915)			
WhiteLaw, C.A., Quackenbush, J., Van Aken, S., Uterback, T.,			
Reenick, A., Frazer, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,			
Citek, R.W., Nuneberg, A., Robbins, D. and Lakey, N.			
Consortium for Maize Genomics			
Unpublished (2002)			
Other GSs: OGIAG08TH			
Contact: Cathy WhiteLaw			
7912 Medical Center Drive, Rockville, MD 20850, USA			
Tel: 301-838-5843			
Fax: 301-838-0208			
Email: whiteLaw@ciagr.org			
Seq primer: TP			
Class: sheared ends.			
Location/Qualifiers			
1. 915			
/organism="Zea mays"			
/mol type="genomic DNA"			
/strain="873"			
/db_xref="taxon:4577"			
/clone="ZMBA0716B15"			
/clone_lib="ZM_0_7_1.5_KB"			
/note="Vector: pGCSK-; Site 1: HincII; 0.7-1.5 kb			
methylation filtered genomic DNA library"			
ORIGIN			
Query Match	51.8%; Score 679; DB 9; Length 915;		
Best Local Similarity	99.9%; Pred. No. 7.7e-132;		
Matches	690; Conservative 0; Mismatches 0; Indels 1; Gaps 1		
621 TGACGTCAGATTTTCTTTTTCATTTCTGTATTTGGTATTTGTTTATATACATTTT	680		
1 TGACGTCAGATTTTCTTTTTC-TCGTGTATTTGTATTTGTTTATATACATTTT	59		
681 CTCTCTTACATAGAGATTTTCTCCGATTTTATTAATGACTAAAGCTATTTT	740		
60 CTCTCTTACATAGAGATTTTCTCCGATTTTATTAATGACTAAAGCTATTTT	119		
741 ATATAGAGAGAGATGCTGATGATCTGCTCAAAATCTTCTGATTTTATTAAGAGC	800		
120 ATATAGAGAGAGATGCTGATGATCTGCTCAAAATCTTCTGATTTTATTAAGAGC	179		
801 TAGTTGGCAACCTGTTCTTTCAAGATTTGATTTTCAAAAAAATAGTTAT	860		
180 TAGTTGGCAACCTGTTCTTTCAAGATTTGATTTTCAAAAAAATAGTTAT	239		
861 TTTCTCTTAAATAGAAACCTTAGAAAAATAGAGTTGGCCAGCTAGCCCTAGAT	920		
240 TTTCTCTTAAATAGAAACCTTAGAAAAATAGAGTTGGCCAGCTAGCCCTAGAT	299		
921 GTTTTCCCAATTAATCAATCACTGTGTATTAATTAATTTGGCCAGCCCATTAATTA	980		
300 GTTTTCCCAATTAATCAATCACTGTGTATTAATTAATTTGGCCAGCCCATTAATTA	359		
981 AAACCGAAATCGAATGAGGAAACCAATCTGAGCTATTTCTTAATTAAGTAAAG	1044		
360 AAACCGAAATCGAATGAGGAAACCAATCTGAGCTATTTCTTAATTAAGTAAAG	419		
1041 GGAAGAGAGAGAAATCAGTTTAAATGATTTGCTCCGAGATGTCGGTTGGCA	1100		
420 GGAAGAGAGAGAAATCAGTTTAAATGATTTGCTCCGAGATGTCGGTTGGCA	479		
1101 CGATAGCCAGCTATCACTAGCTCATAGGTGCTAGCTCAGGTTCCGAGCTCTGTGTC	1166		
480 CGATAGCCAGCTATCACTAGCTCATAGGTGCTAGCTCAGGTTCCGAGCTCTGTGTC	539		
1161 ATTCACATGGCATATCAATGCTGTTCAACCGTGTCTGTTCCATCGCCAAAGCT	1222		
540 ATTCACATGGCATATCAATGCTGTTCAACCGTGTCTGTTCCATCGCCAAAGCT	599		

Oy	1221	TGCCATCTCTGAACCAAGAGATACCTACTGCCAAACATCTACTTCAATCACTT	1280
Db	600	TGCCATCTCTGAACCAAGAGATACCTACTGCCAAACATCTACTTCAATCACTT	659
Oy	1281	CCATGCCAACACGCGCATATGTCTTCCCTAAC	1311
Db	660	CCATGCCAACACGCGCATATGTCTTCCCTAAC	690

RESULT 3					
CC656939/c					
LOCUS	CC656939	687 bp	DNA	linear	GSS 19-JUN-2003
DEFINITION	CGMDQ20TV ZM_0.7_1.5_KB Zee maye genomic clone ZMMBMA0554D15,				

ACCESSION	CC656939	GI:32060231
VERSION	CC656939.1	
KEYWORDS	GS.	
SOURCE	Zea mays	
ORGANISM	Zea mays	

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 687)	Whitelaw C.A., Quackenbush, J., Van Aken, S., Uterback, T., Reenick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Cleek, R.W., Nurnberg, A., Robbins, D. and Lakey, N.	Consortium for Maize Genomics	Unpublished (2002)	Other GSSs: OGDQ207M

9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: white@wglr.org
Seq primer: TF
Class: sheared ends.

```

FEATURES
SOURCE
location/Qualifiers
1..687
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM8BMA055d15"
/clone_1lb="ZM_0.7_1.5_KB"
/note="Vector: pGCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

```

Query Match	32.0%	Score 419.8	DB 9	Length 687
Beats Local Similarly	98.5%	Pred. No. 1.6e-77		
Matches 421, Conservative	0	Mismatches 2	Indels 0	Gaps 0

Qy	889	GAAAAATAGAGTTGCGACACTAGAGCCCTGGAATGTTTTCCCAATAAATTACAACTACGTGTG	948
Db	687	GAAAAATAGAGTGCCCACTAGAGCCCTGGAATGTTTTCCCAATAAATTACAACTACGTGTG	628
Qy	949	TATATATTTTGGCGACGCCCATTAATTTATTTAAACCGAACTGAAATCGAGCGAAACCA	1008
Db	627	TATATATTTTGGCCAGGCCCATTAATTTATTTAAACCGAACTGAAATCGAGCGAAACCA	568
Qy	1009	AATCTGAGCTATTTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAATCAAGTTTAA	1068
Db	567	AATCTGAGCTATTTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAATCAAGTTTAA	508
Qy	1069	AGTCATGTCCCTGAGATGTGGGGTTTGGCAACGATAGCCACCCGTAATCATAGCTCATAG	1128
Db	507	AGTCATGTTCCTTGAGATGTGGGGTTTGGCAACGATAGCCACCCGTAATCATAGCTCATAG	448
Qy	1129	GTGCTTACGTCAAGTTTGGCGAGCTCTGCTGTATCTCAATGGCAATCAATCATGCTGTTT	1188
Db	447	GTGCTTACGTCAAGTTTGGCGAGCTCTGCTGTATCTCAATGGCAATCAATCATGCTGTTT	388

Qy	Db	Qy	Db
1189	387	1249	327
CAACCGTTCGCTGTTCATTCATCGCCAAAGCTTGCCATTTCTGAAACAAGATACCTA	CAACCGTTCGCTGTTCATTCATCGCCAAAGCTTGCCATTTCTGAAACAAGATACCTA	CTCCCAACCAATCATTTACTCATGCAACTTCAGAGCAACAGGCAATATGTTCTCG	CTCCCAACCAATCATTTACTCATGCAACTTCAGAGCAACAGGCAATATGTTCTCG
1248	328	1308	268

Oy	1309 AAC 1311
Db	267 AAC 265

RESULT 4	
CG52571/c	715 bp DNA linear GSS 25-AUG-2003
LOCUS	
DEFINITION	CG52571 CG48B0572M 0_7_1_5_KB Zea mays genomic clone ZmBBA0809B10, genomic survey sequence.
ACCESSION	CG52571
VERSION	CG52571.1 GI:34154661
KEYWORDS	GSS.
SOURCE	Zea mays
ORGANISM	Zea mays

REFERENCE AUTHORS	TITLE JOURNAL COMMENT
1 (bases 1 to 715) Whiteclaw C.A., Quackenbush, J., Van Aken, S., Utecherak, T., Resnick A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.	Consortium for Maize Genomics Unpublished (2002) Contact: Cathy Whiteclaw

9112 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Tel: 301-838-5843
Fax: 301-838-0208
Email: white.law@tigr.org
Seq primer: TP
Class: sheared ends.
Location/Qualifiers
1. .715

```

FEATURES
    source
        Location/Qualifiers
            1..715
                /organism="Zea mays"
                /mol_type="genomic DNA"
                /strain="B73"
                /db_xref="taxon:4577"
                /clone="ZM8BMA0809810"
                /clone_1kb="ZM_0.7.1.5_KB"
                /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

```

```

Query Match      7.3%; Score 96; DB 9; Length 715;
Best Local Similarity 80.1%; Pred. No. 1.2e-09;
Matches 125; Conservative 0; Mismatches 30; Indels 1; Gaps 1;

```

Qy	135	TCCTCCACCCGCGGTTCACATTTCTTGGANATCGGAGTATCCCATCTGACCGAGGCCCA	194
Db	356	TCACACACCCGCGCTTTGTGCTTTCTTGGATGTGGCGGTTCCCGCTGATCGAGGCCCA	297
Qy	195	TCAGACACCTTTGGGACACCCCATCAAGGCGCTTTGGATGGCCCAACGATCGGG	254
Db	296	TAAAGGACCTGTCCGAATCCCATTTAAAGCTTTCCGAATGGCCCAAGGACAT-CG99G	238
Qy	255	TCGTGTGATCCAGGGGATATATGTCTCCCAATC	290
Db	237	TCGTGTGACCCGGGGATATCTATCCCCACAAC	202

RESULT 5	LOCUS	DEFINITION
CL235046	CL235046	967 bp DNA linear GSS 15-JAN-2004
	ZMMBB0575001r	ZMMBB (HindIII) Zea mays genomic clone

ACCESSION	ZMMBB0575001.3', genomic survey sequence.
VERSION	CL235046
KEYWORDS	CL235046.1 GI:40891729
SOURCE	GSS.
ORGANISM	Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD clade; Panicoideae; Andropogoneae; Zea.	
1 (bases 1 to 967)	
Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C., Zohovetz,V., Fuks,G., Yu,Y., Wing,R. and Messing,J.	
Sequencing of the maize genome at PGIR (2003c)	
Unpublished (2003)	
Contact: Bharti,A.K.	
Dr.Joachim Messing's lab	
The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers University	
190 Fellinghuysen Road, Piscataway, NJ 08854, USA	
Tel: 732 445 3801	
Fax: 732 445 5735	
Email: bharti@waksman.rutgers.edu	
Seq primer: SP6	
Class: BAC ends	
High quality sequence start: 64.	
Location/Qualifiers	
1..967	
/organism="Zea mays"	
/mol_type="genomic DNA"	
/cultivar="B73"	
/db_xref="taxon:4577"	
/clone="ZMMBB0575001"	
/lab_host="E. coli DH10B"	
/clone_lib="ZMMBB (HindIII)"	
/note="Vector: pCUG1; Site_1: HindIII; Site_2: HindIII"	
ORIGIN	
Query Match	7.0%; Score 92; DB 9; Length 967;
Best Local Similarity	78.6%; Pred. No. 8.5e-09;
Matches 110; Conservative 0; Mismatches 30; Indels 0; Gaps 0;	
Oy	150 TGCACATTTCTTGATGTGCGGTGTTCCCATCTGACGAGGCGCCATCAGACCTTTGG 209
Db	827 TGGCGCTTTCTTGATACCGCGGTTCCCGTCTATTAGAGCCATCCAGACCTTTGG 886
Oy	210 GACACCCATCAAGGGCTTTGGATGCGCCACGAGACGTATCGGCTGTTGATTCAG 269
Db	887 GTGGCGCCATCAAGTGCTTTGGATGCTCACGAGGATGATCGGCGCGTGTGACTGG 946
Oy	270 GGATATATGTCCTCCCAAT 289
Db	947 GGATATCTGTCCCAATAGT 966
RESULT 6	
CG048704	814 bp DNA linear GSS 19-AUG-2003
LOCUS	PULL019TB ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa0611c13,
DEFINITION	genomic survey sequence.
ACCESSION	CG048704
VERSION	CG048704.1 GI:33920884
KEYWORDS	GSS.
SOURCE	Zea mays
Zea mays	
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD clade; Panicoideae; Andropogoneae; Zea.	
1 (bases 1 to 814)	
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utechtack,T., Benneick,A., Frazer,C.M., Yuan,Y., San Miguel,P., Ma,J. and	
Reinisch,A.J.	
Maize Genomics Consortium	
Unpublished (2003)	
TITLE	
JOURNAL	

COMMENT	Other GSSs: PUIL019TD
Contact: Cathy Whitelaw	
TIGR	
9712 Medical Center Drive, Rockville, MD 20850, USA	
Tel: 301-838-5843	
Fax: 301-838-0208	
Email: whitelaw@cigr.org	
Seq primer: TR	
Class: sheared ends.	
FEATURES	Location/Qualifiers
Source	1..814
	/organism="Zea mays"
	/mol_type="genomic DNA"
	/strain="B73"
	/db_xref="taxon:4577"
	/clone="ZMMB0611C13"
	/clone_1lb="ZM_0.6-1.0_KB"
	/note="Vector: pCR4-TOP0; Site_1: EcoRI; 0.6-1.0 kb high
	Cot selected genomic DNA library"
ORIGIN	
Query Match	6.7%; Score 88; DB 9; Length 814;
Best Local Similarity	62.7%; Pred. No. 5.8e-08;
Matches 153; Conservative	0; Mismatches 90; Indels 1; Gaps 1;
Db	713 TTATATAAATGACTATTAAGTCAATTTTATATTAAGACGACGATGCTGATCTCGTT 772
Qy	
Db	487 TTATAGAGACGCTGTGTAGACATGGAAGATATAGAAAAATATCTTTAGAGAAATGTTA 546
Qy	773 CAAAAATCTTCGATTTTATTAAGACTAGTTGGCAACCCGTCTTCTTCAAGAAAT 832
Db	547 TAAAGACAGAAACATTTCTTAAGAGCTAGTTGGTAATCCCATTTTTCAAAGATT 606
Qy	833 TTGATTTTTCAAAA-AAAATAGTTATTTCTCTTTATAAATAGAAAACACTTAGAA 891
Db	607 TTCATTCTCAAGAGGAAAAATGTTATTTCCCTTTTGTAAAAATGGTATATCTGGAA 666
Qy	892 AAATAGATTGCCAGACTAGCCCTAGAAATGTTTCCCAATAAATTAACAATCACTGTAT 951
Db	667 AAATAGATTCTCAAACTAGCCCTAAGATTATATCCGATTAATCCCTATCTATCTCT 726
Qy	952 AATT 955
Db	727 ATTT 730
RESULT 7	
CG414922/c	754 bp DNA 1linear GSS 08-SEP-2003
LOCUS	ZMMBB0250L09.r ZMMBB Zea mays genomic clone ZMMBB0250L09 3',
DEFINITION	genomic survey sequence.
ACCESSION	CG414922
VERSION	CG414922.1 GI:34505144
KEYWORDS	GSS.
SOURCE	
ORGANISM	Zea mays
	Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE	1. (bases 1 to 754)
AUTHORS	Yu, Y., Kim, H.R., Hatfield, J., Soderlund, C., Bharti, A.K., Messing, J. and Wing, R.
TITLE	Sequencing of the maize genome
JOURNAL	Unpublished (2003)
COMMENT	Contact: Rod Wing Arizona Genomics Institute University of Arizona Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ 85721-0088 USA Tel: 520 626 3967 Fax: 520 621 9288 Email: http://genome.arizona.edu
	PCR Primers

Db 461 TATTTTTCAGAGTAATTAAGTTTATTTTCATGCGAATAAGAAATTAATGGAAAA 520
QY 894 ATAGATGGCAGACTAGCCCTAGA 918
Db 521 ATTAGTTTCCAAACTAGCCCTAAA 545

RESULT 10
CC400575/c 797 bp DNA linear GSS 19-MAY-2003
LOCUS PNHU617D_ZM_0.6_1.0_KB_Zea_mays_genomic_clone_ZM8BTA480L01,
DEFINITION genomic survey sequence.

ACCESSION CC400575
VERSION CC400575.1 GI:30880665
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 797)

AUTHORS WhiteLaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.

TITLE Maize Genomics Consortium
JOURNAL Unpublished (2003)
COMMENT Other GSSs: PNHU617B
Contact: Cathy WhiteLaw
TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208

Email: whiteLaw@tigr.org
Seq primer: TP

Class: sheared ends.

FEATURES Location/Qualifiers

source

1..797
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM8BTA480L01"
/clone_lib="ZM_0.6_1.0_KB"
/note="Vector: PCR4-TOPO; Site: 1: EcoRI; 0.6-1.0 kb high
COT selected genomic DNA library"

ORIGIN

Query Match 6.6%; Score 86.6; DB 8; Length 797;
Best Local Similarity 59.7%; Pred. No. 1.1e-07;

Matches 163; Conservative 0; Mismatches 109; Indels 1; Gaps 1;

QY 713 TTTATTAATGACTATTAAGTCAATTTTATATAGAGCAGCATGTCGTAGATTCGTT 772
Db 577 TTTAGGAGACGTTGGTAGAGATGAAAGATATAGAAAATATCTTTTAGAGATGTTA 518
QY 773 CAAAAATCTTCTGATTTTATAGAGCTAGTTGGCAACCTGTTCTTTCAAAGATT 832
Db 517 TAAAGACAGAAAACATTTTATAGAGCTAGTTGGTAATCCCATTTTTCAAAGATT 458
QY 833 TTGATTTTTCAAA-AAAATTAAGTTATTTTCTTTATAAATGAAAACACTTAGAA 891
Db 457 TTCAATCTCAAGAGGAAAATAGTTATTTCCCTTTGAAAATGATATCATCTCGGAA 398
QY 892 AATTAAGTTGGCAGACTAGCCCTAGAAATGTTTCCCAATTAATTAATCAATCACTGTAT 951
Db 397 AATTAAGTTCTCAAACTAGCCCTTAGAGTTATATCCGATTAATCCCTATCTCATCTCT 338
QY 952 AATTATTTGGCCAGCCCATTAATTAATTTTAAAC 984
Db 337 ATTTCAAATTCACCTTGGCAAAATAGTGTATC 305

RESULT 11

CC400574 820 bp DNA linear GSS 19-MAY-2003
LOCUS PNHU617B_ZM_0.6_1.0_KB_Zea_mays_genomic_clone_ZM8BTA480L01,
DEFINITION genomic survey sequence.

ACCESSION CC400574
VERSION CC400574.1 GI:30880664
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 820)

AUTHORS WhiteLaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.

TITLE Maize Genomics Consortium
JOURNAL Unpublished (2003)
COMMENT Other GSSs: PNHU617D
Contact: Cathy WhiteLaw
TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208

Email: whiteLaw@tigr.org
Seq primer: TP

Class: sheared ends.

FEATURES Location/Qualifiers

source

1..820
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM8BTA480L01"
/clone_lib="ZM_0.6_1.0_KB"
/note="Vector: PCR4-TOPO; Site: 1: EcoRI; 0.6-1.0 kb high
COT selected genomic DNA library"

ORIGIN

Query Match 6.6%; Score 86.6; DB 8; Length 820;
Best Local Similarity 59.7%; Pred. No. 1.1e-07;

Matches 163; Conservative 0; Mismatches 109; Indels 1; Gaps 1;

QY 713 TTTATTAATGACTATTAAGTCAATTTTATATAGAGCAGCATGTCGTAGATTCGTT 772
Db 374 TTTAGGAGACGTTGGTAGAGATGAAAGATATAGAAAATATCTTTTAGAGATGTTA 433
QY 773 CAAAAATCTTCTGATTTTATAGAGCTAGTTGGCAACCTGTTCTTTCAAAGATT 832
Db 434 TAAAGACAGAAAACATTTTATAGAGCTAGTTGGTAATCCCATTTTTCAAAGATT 493
QY 833 TTGATTTTTCAAA-AAAATTAAGTTATTTTCTTTATAAATGAAAACACTTAGAA 891
Db 494 TTCAATCTCAAGAGGAAAATAGTTATTTCCCTTTGAAAATGATATCATCTCGGAA 553
QY 892 AATTAAGTTGGCAGACTAGCCCTAGAAATGTTTCCCAATTAATTAATCAATCACTGTAT 951
Db 554 AATTAAGTTCTCAAACTAGCCCTTAGAGTTATATCCGATTAATCCCTATCTCATCTCT 613
QY 952 AATTATTTGGCCAGCCCATTAATTAATTTTAAAC 984
Db 614 ATTTCAAATTCACCTTGGCAAAATAGTGTATC 646
RESULT 12
CG103452 471 bp DNA linear GSS 20-AUG-2003
LOCUS PUIBE19TB_ZM_0.6_1.0_KB_Zea_mays_genomic_clone_ZM8BTA628D13,
DEFINITION genomic survey sequence.
ACCESSION CG103452
VERSION CG103452.1 GI:33985746
KEYWORDS GSS.
SOURCE Zea mays

ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 471)
AUTHORS WhiteJaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
TITLE Maize Genomics Consortium
JOURNAL Unpublished (2003)
COMMENT Other GSSs: PUBB19TD
Contact: Cathy WhiteJaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteJaw@tigr.org
Seq primer: TR
Class: sheared ends.

FEATURES
source Location/Qualifiers
1..471
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZM_0.6-1.0_KB"
/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
cot selected genomic DNA library"

ORIGIN
Query Match 6.5%; Score 85.4; DB 9; Length 471;
Best Local Similarity 69.5%; Pred. No. 2e-07;
Matches 116; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

LOCUS 789 TTTTAAAGAGCTAGTTGGCAACCCGTTCTTTCAAAATTTGATTTTCAAAA 848
DEFINITION 216 TTTTCTTAAGAGCTAGTTGAAATCCCATTTTCCAGGATTTACATTTTTCAGGG 275
ACCESSION 849 AATTAAGTTATTTCTCTTATTAATAAGAAAAACCTTAGAAAAATAGATTGCCAGC 908
VERSION 276 AATTAAGTTATTTCTCTTATTAATAAGAAAAACCTTAGAAAAATAGATTGCCAGC 335
KEYWORDS Db
SOURCE 909 TAGCCCTAGATGTTTCCCAATTAATTAACAATCACTGTGTATTAATT 955
ORGANISM 336 TAGCCCTAGATGTTTCTTAATAAGAAAAATAGATCAATGAATAGATT 382

RESULT 13
CG082135 765 bp DNA linear GSS 20-AUG-2003
LOCUS PUFQ12TD_ZM_0.6-1.0_KB Zea mays genomic clone ZM8BTA0712B24,
DEFINITION genomic survey sequence.
ACCESSION CG082135
VERSION CG082135.1 GI:33964429
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 765)
AUTHORS WhiteJaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
TITLE Maize Genomics Consortium
JOURNAL Unpublished (2003)
COMMENT Other GSSs: PUFQX12TB
Contact: Cathy WhiteJaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208

Email: whiteJaw@tigr.org
Seq primer: TR
Class: sheared ends.
Location/Qualifiers
1..765
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZM_0.6-1.0_KB"
/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
cot selected genomic DNA library"

ORIGIN
Query Match 6.5%; Score 85.2; DB 9; Length 765;
Best Local Similarity 71.8%; Pred. No. 2e-07;
Matches 125; Conservative 0; Mismatches 48; Indels 1; Gaps 1;

LOCUS 787 ATTTTAAAGAGCTAGTTGGCAACCCGTTCTTTCAAAATTTGATTTTCAAA 846
DEFINITION 255 ATCTTCTAATGACCTAGTTGGAAACCTTATTT-TCACAGGATTTTATTTTCAAT 197
ACCESSION 847 AATTAAGTTATTTCTCTTATTAATAAGAAAAACCTTAGAAAAATAGATTGCCAG 906
VERSION 196 GAAATTAAGTTATTTCTCTTATTAATAAGAAAAACCTTAGAAAAATAGATTGCCAG 137
KEYWORDS Db
SOURCE 907 ACTAGCCCTAGATGTTTCCCAATTAATTAACAATCACTGTGTATTAATTG 960
ORGANISM 136 ACTAGCCCTAAGATTTTATTAATAAGAAAAATAGATGCTGATTTTCAATTG 83

RESULT 14
CC630219 781 bp DNA linear GSS 19-JUN-2003
LOCUS OGUG53TV_ZM_0.7-1.5_KB Zea mays genomic clone ZMMEMa0406J09,
DEFINITION genomic survey sequence.
ACCESSION CC630219
VERSION CC630219.1 GI:32003354
KEYWORDS Db
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 781)
AUTHORS WhiteJaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Resnick,A., Fraser,C.M., Buddiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
TITLE Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Other GSSs: OGUG53TV
Contact: Cathy WhiteJaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteJaw@tigr.org
Seq primer: TR
Class: sheared ends.

FEATURES
source Location/Qualifiers
1..781
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZMMEMa0406J09"
/note="Vector: pBCSK-1; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match 6.5%; Score 85.2; DB 9; Length 781;

Search completed: September 16, 2005, 08:08:42
Job time : 4089.83 secs

Best Local Similarity 71.8%; Pred. No. 2.2e-07;
Matches 125; Conservative 0; Mismatches 48; Indels 1; Gaps 1;

```
QY 787 ATTTTAAAGACGAGTTGGCAACCGTCTTCTTCAAGAATTTGATTTTCAAA 846
    |||||
DB 665 ATCTTTCTAATGACTAGTTGGAGACCTTATTT-TCTCAGGGATTTTATTTTCAAT 607
    |||||
QY 847 AAAAATTAGTTTCTTCTTATATAAATAGAAAACCTTAGAAAATAGAGTTGCCAG 906
    |||||
DB 606 GAAAAATTAGTTTATTTCTTCTTGAGAAAATATTAATCACTTGAGAAAATATAGTTTCCA 547
    |||||
QY 907 ACTAGCCCTAGATGTTTCCCAATTAATACATCATGCTGATATATTATTG 960
    |||||
DB 546 ACTAGCCCTAAGATTTTATATAAATAGAGAAAATGCTCTATTTTCATTG 493
    |||||
```

RESULT 15

CG349565 815 bp DNA linear GSS 26-AUG-2003
LOCUS OG0FI83TH_ZM_0_7_1.5_Zea_mays_genomic_clone_ZMM8Ma0705N21,
DEFINITION genomic survey sequence.

ACCESSION CG349565
VERSION CG349565.1 GI:34266831
KEYWORDS GSS.

SOURCE

ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 815)

REFERENCE

AUTHORS White, C.A., Quackenbush, J., Van Aken, S., Uteerback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robb, D. and Lake, N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSs: OG0FI83TV
Contact: Cathy White, law
TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitec@tigr.org

Seq primer: TR
Class: sheared ends.

FEATURES

Location/Qualifiers

1..815
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_11b="ZM_0.7_1.5_KB"
/note="Vector: pBCK-; Site_1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match 6.5%; Score 85.2; DB 9; Length 815;
Best Local Similarity 71.8%; Pred. No. 2.2e-07;

Matches 125; Conservative 0; Mismatches 48; Indels 1; Gaps 1;

```
QY 787 ATTTTAAAGACGAGTTGGCAACCGTCTTCTTCAAGAATTTGATTTTCAAA 846
    |||||
DB 173 ATCTTTCTAATGACTAGTTGGAGACCTTATTT-TCTCAGGGATTTTATTTTCAAT 115
    |||||
QY 847 AAAAATTAGTTTCTTCTTATATAAATAGAAAACCTTAGAAAATAGAGTTGCCAG 906
    |||||
DB 114 GAAAAATTAGTTTATTTCTTCTTGAGAAAATATTAATCACTTGAGAAAATATAGTTTCCA 55
    |||||
QY 907 ACTAGCCCTAGATGTTTCCCAATTAATACATCATGCTGATATATTATTG 960
    |||||
DB 54 ACTAGCCCTAAGATTTTATATAAATAGAGAAAATGCTCTATTTTCATTG 1
    |||||
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2005, 08:14:25 ; Search time 4104 Seconds

(without alignments)
16458.716 Million cell updates/sec

Title: US-10-713-381-1

Perfect score: 1394
Sequence: 1 ccatggtgtctctatgaa.....cttttcgtccaccaccatg 1394

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

GenBml:*
1: gb_ba:*
2: gb_bcg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_stb:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1394	100.0	1394	6	AX224394 Sequence
2	1394	100.0	1394	6	BD062176 Male t18
3	1389	99.6	1394	8	AF360356 Zea mays
4	1311	94.0	1394	6	AX224395 Sequence
5	1311	94.0	1394	6	BD062177 Male t18
6	157	11.3	255	6	AX224402 Sequence
7	111	8.0	158	6	AX224396 Sequence
8	50	3.6	50	6	AX224398 Sequence
9	40	2.9	40	6	AX224399 Sequence
10	30	2.2	30	6	AX224397 Sequence
11	25	1.8	114817	2	AC149836 Zea mays
12	24	1.7	10384	1	AE010607 Fusobacte
13	24	1.7	184509	2	AC113237 Canis fam
14	24	1.7	212621	2	AC098210 Rattus no
15	24	1.7	213753	2	AC110715 Rattus no
16	24	1.7	239424	2	AC118504 Rattus no
17	24	1.7	245072	2	CR381551 Danio rer
18	24	1.7	252679	2	AC118394 Rattus no
19	23	1.6	849	6	AX384588 Sequence

20	23	1.6	93615	9	AC004950 Homo sapi
21	23	1.6	106769	9	AL357079 Human DNA
22	23	1.6	114214	8	AC125479 Medicago
23	23	1.6	120180	5	BX323592 Zebrafish
24	23	1.6	132698	2	AC005051 Homo sapi
25	23	1.6	164304	2	AP002398 Homo sapi
26	23	1.6	159876	10	AC124433 Mus muscu
27	23	1.6	179676	9	AC027514 Homo sapi
28	23	1.6	187002	2	AC134483 Rattus no
29	23	1.6	204724	10	AC131912 Mus muscu
30	23	1.6	206924	10	AC087780 Mus muscu
31	23	1.6	228736	2	AC121614 Rattus no
32	23	1.6	245121	2	AC126482 Rattus no
33	23	1.6	255925	2	AC094360 Rattus no
34	23	1.6	259743	2	AC113965 Mus muscu
35	22	1.6	939	3	CG1579915 Crasosotr
36	22	1.6	1154	3	AY617812 Sterkell
37	22	1.6	28168	3	CEY3799A Al032625 Caenorhab
38	22	1.6	93132	10	AC084402 Mus muscu
39	22	1.6	128625	8	AC147499 Medicago
40	22	1.6	128907	9	AC105445 Homo sapi
41	22	1.6	135726	8	AC147498 Medicago
42	22	1.6	137477	5	BX004882 Zebrafish
43	22	1.6	158833	8	AC134926 Oryza sat
44	22	1.6	170059	9	AC026780 Homo sapi
45	22	1.6	172117	9	AC024890 Homo sapi

ALIGNMENTS

RESULT 1	AX224394	1394 bp	DNA	linear	PAT 10-SEP-2001
LOCUS	Sequence 1	from Patent WO0160997.			
DEFINITION	AX224394				
ACCESSION	AX224394.1	GI:15554636			
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
	Zea mays				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD				
	clade; Panicoideae; Andropogoneae; Zea.				
REFERENCE					
AUTHORS	1	Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.			
TITLE		Male tissue-preferred regulatory region and method of using same			
JOURNAL		Patent: WO 0160997-A 1 23-AUG-2001;			
		PIONEER HI-BRD INTERNATIONAL, INC. (US)			
FEATURES					
source					
	1..1394				
	/organism="Zea mays"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:4577"				
ORIGIN					
Query Match	100.0%;	Score 1394;	DB 6;	Length 1394;	
Best Local Similarity	100.0%;	Pred. No. 0;			
Matches 1394;	Conservative	0;	Mismatches	0;	Indels 0;
					Gaps 0;
QY	1	CCATGCTGCTCTATGAAAAGATGATGATGATGCTATGCTTTCCTTTTAAAGGTC	60		
	1	CCATGCTGCTCTATGAAAAGATGATGATGATGCTATGCTTTCCTTTTAAAGGTC	60		
DB	61	CTTCTTCCTCTATTACTGATCGATCGGGTTACAAAACCTTCCACGGGTGCATGAT	120		
	61	CTTCTTCCTCTATTACTGATCGATCGGGTTACAAAACCTTCCACGGGTGCATGAT	120		
QY	121	CTTCATGTTCCATCTTCCACCTCGGTTGCACTTTTGGATGTCGTTCCAT	180		
	121	CTTCATGTTCCATCTTCCACCTCGGTTGCACTTTTGGATGTCGTTCCAT	180		
DB	121	CTTCATGTTCCATCTTCCACCTCGGTTGCACTTTTGGATGTCGTTCCAT	180		
	181	CTGACGAGCCCATGACACCTTTCGGGACACCATCAAGGCTTTCGATGGCCCA	240		
	181	CTGACGAGCCCATGACACCTTTCGGGACACCATCAAGGCTTTCGATGGCCCA	240		

QY 361 GGGGCTGAGATGATTTGGCTTAGGGCCCCGAGAAATGCGAGACCGACCATGTCTAGTG 420
DB 361 GGGGCTGAGATGATTTGGCTTAGGGCCCCGAGAAATGCGAGACCGACCATGTCTAGTG 420
QY 421 TCCACTATTGGCACTACCCAGAACAAAGATTAAAAAATAACCAAGTAATCAATCCACT 480
DB 421 TCCACTATTGGCACTACCCAGAACAAAGATTAAAAAATAACCAAGTAATCAATCCACT 480
QY 481 CGAAAGCTATCATGTAAATGTTTAAAGAAACATCTATTAAAAACGATCCCTTTAAAAAA 540
DB 481 CGAAAGCTATCATGTAAATGTTTAAAGAAACATCTATTAAAAACGATCCCTTTAAAAAA 540
QY 541 CAAGCAATTTCCAGAAAGACAAATTTGTTACAGTTTCAAAACATCTAAGACGACAA 600
DB 541 CAAGCAATTTCCAGAAAGACAAATTTGTTACAGTTTCAAAACATCTAAGACGACAA 600
QY 601 TTATATCGAAAGGTAGCTATGACGTTCAAGATTTTCTTTTCAATCTGTTATTTTGT 660
DB 601 TTATATCGAAAGGTAGCTATGACGTTCAAGATTTTCTTTTCAATCTGTTATTTTGT 660
QY 661 ATTGTTTTTAAATACATTTTCTTCTTACAAATAGAGTGAATTTTCCGATTTTAA 720
DB 661 ATTGTTTTTAAATACATTTTCTTCTTACAAATAGAGTGAATTTTCCGATTTTAA 720
QY 721 ATGACTATAAGCTATTTTATATAGAGACGCGATGCTAGATTTCTCGTCAAAATC 780
DB 721 ATGACTATAAGCTATTTTATATAGAGACGCGATGCTAGATTTCTCGTCAAAATC 780
QY 781 TTTCTGATTTTAAAGAGCTAGTTTGGCAACCTGTTTCTTTCAAAAGATTTTGAATTT 840
DB 781 TTTCTGATTTTAAAGAGCTAGTTTGGCAACCTGTTTCTTTCAAAAGATTTTGAATTT 840
QY 841 TTCAAAAAAATAGTTTATTTCTTTTAAATAAGAAACATTTAGAAAAATAGAGT 900
DB 841 TTCAAAAAAATAGTTTATTTCTTTTAAATAAGAAACATTTAGAAAAATAGAGT 900
QY 901 TGGCAGACTAGCCCTAGAAATGTTTCCCATTAATTCATCTGCTATTAATTTTGG 960
DB 901 TGGCAGACTAGCCCTAGAAATGTTTCCCATTAATTCATCTGCTATTAATTTTGG 960
QY 961 GCCAGCCCCATTAATTTTAAACCGAACTGAATTCGACGAAACCAATCTGAGCTAT 1020
DB 961 GCCAGCCCCATTAATTTTAAACCGAACTGAATTCGACGAAACCAATCTGAGCTAT 1020
QY 1021 TTTCTAGATTAGTAAAGAGAGAGAGAGAGAGAGAAATCAGTTTAACTCATTTGCC 1080
DB 1021 TTTCTAGATTAGTAAAGAGAGAGAGAGAGAGAGAAATCAGTTTAACTCATTTGCC 1080
QY 1081 TGAAGATGTGGGTTTGGCAACGATAGCCACCTTATATAGCTCATAGTGCTTACGTCA 1140
DB 1081 TGAAGATGTGGGTTTGGCAACGATAGCCACCTTATATAGCTCATAGTGCTTACGTCA 1140
QY 1141 GGTTCGGCAGACTCTGCTGATCTCAGATGCAATCTACATGCTTGTTCACACCTTGCTC 1200
DB 1141 GGTTCGGCAGACTCTGCTGATCTCAGATGCAATCTACATGCTTGTTCACACCTTGCTC 1200
QY 1201 TTGTTTCATGTCGCAAGCCTTGCTTATCTGAAACCAAGAGATACCTATCCCAACAT 1260
DB 1201 TTGTTTCATGTCGCAAGCCTTGCTTATCTGAAACCAAGAGATACCTATCCCAACAT 1260
QY 1261 CCATTTTACATGCAACTTCCATGCAACGACCATATGTTTCTGCAACCAATCATAT 1320
DB 1261 CCATTTTACATGCAACTTCCATGCAACGACCATATGTTTCTGCAACCAATCATAT 1320
QY 1321 AAAGATCAACAAGACTAGCTTCTCCGCTAGCTTCCCTCTCTCTCTGCGCATTTT 1380
DB 1321 AAAGATCAACAAGACTAGCTTCTCCGCTAGCTTCCCTCTCTCTCTGCGCATTTT 1380
QY 1381 CGTCCACCAACCATG 1394
DB 1381 CGTCCACCAACCATG 1394
QY 1394 CGTCCACCAACCATG 1394
DB 1394 CGTCCACCAACCATG 1394

RESULT 3
AF360356 3343 bp DNA linear pln 12-MAY-2001
LOCUS Zea mays male fertility protein (M645) gene, complete cds.
DEFINITION
ACCESSION AF360356
VERSION AF360356.1 GI:114028756
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 3343)
Fox,T.W., Trimmell,M.R. and Albertsen,M.C.
AUTHORS Cloning of M645, a gene required for male fertility from Zea mays
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 3343)
Fox,T.W., Trimmell,M.R. and Albertsen,M.C.
AUTHORS Direct Submission
TITLE Submitted (13-MAR-2001) Trait and Technology Development, Pioneer
H1-Bred Incl. Inc., 7300 N.W. 62nd Ave., P.O. Box 1004, Johnston,
IA 50131-1004, USA
FEATURES
source
1..3343
/organism="Zea mays"
/mol_type="genomic DNA"
/culivar="B73"
/db_xref="taxon:4577"
/chromosome="9L"
/c1392..>2942
/gene="M645"
/join(<1392..1768,1898..2182,2280..2447,2534..>2942)
/product="male fertility protein"
/join(1392..1768,1898..2182,2280..2447,2534..2942)
/gene="M645"
/codon_start=1
/product="male fertility protein"
/protein_id="AAK52489.1"
/translation="MEKNILQWRGRDGIVOYPHLFPALALALVADPFGSLAEV
DYRPVKHELAPYGEVMSGWPNDASRLRGRLPEYGVGPESTIRPDLOGRPGVGLA
DERVVRMGESEAGHETPAVMPNDSEBEYCAKCVNSTTRKHKEKEFCGRPLRPHGE
TSELVADAVYGLMNVGSGVASSVAEADSDPIRFANDLDVHNGSVFTDTSNRX
SRKDHNTLLEBEGTGLRLRYDPEYSGVNVYKGLVFPNGVDIISDHQFLRSETTNC
RIMRYMLBEPRAEVEFANLPFPDNRVSNRQFPWALDCRTPADQVFAKRPMLR
TLVFPRLSLKYLTKAARRMHTVALLDSEGRVVEVLEDRHEHMKLVSEVREVGK
LMIGVANHNTIIPYLED"

ORIGIN
Query Match 99.6%; Score 1389; DB 8; Length 3343;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1389; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCATGCTGCTCTATGAAAAAGATGACATGCTCTATATCCGTTTCTTAGGGTCC 60
DB 1 CCATGCTGCTCTATGAAAAAGATGACATGCTCTATATCCGTTTCTTAGGGTCC 60
QY 61 CTTCTTCGCTTATTAACGACGTAATGCGGGTTCAAAAACTTCCACGGGTGACATGAT 120
DB 61 CTTCTTCGCTTATTAACGACGTAATGCGGGTTCAAAAACTTCCACGGGTGACATGAT 120
QY 121 CTCGATGTTCACTTCTCCACCTCGCTTGACATTTCTTGATGTCGCTGCTCCAT 180
DB 121 CTCGATGTTCACTTCTCCACCTCGCTTGACATTTCTTGATGTCGCTGCTCCAT 180
QY 181 CTGACCGAGGCCCATCAACACCTTTGGGAAACCCATCAAGGGCTTTCCGATGGCCCA 240
DB 181 CTGACCGAGGCCCATCAACACCTTTGGGAAACCCATCAAGGGCTTTCCGATGGCCCA 240
QY 241 CGAGAGTATCGGGTGTGTGATCCAGGGGATATATGTCCTCCCAATCCGACCTATA 300
DB 241 CGAGAGTATCGGGTGTGTGATCCAGGGGATATATGTCCTCCCAATCCGACCTATA 300

Db 241 CGAGAGCTATCGGGTGTGGATCCAGGGATATATGTCGCCACAAATGTCACCTATA 300
 Qy 301 TTATTTATCTTTAGATTTTATTTATTTTGGAAAAATACAACTTATCTTTGTGTA 360
 Db 301 TTATTTATCTTTAGATTTTATTTATTTTGGAAAAATACAACTTATCTTTGTGTA 360
 Qy 361 GGGCTCAGCATGATTTTGGCTTAAAGGCGCCAGAAATGCGAGACCGCATGCTAGTG 420
 Db 361 GGGCTCAGCATGATTTTGGCTTAAAGGCGCCAGAAATGCGAGACCGCATGCTAGTG 420
 Qy 421 TCCACTATTGGCACTACCCAGAAACAAGATTTAAAAAATAACAAAGTAACTAATCCACT 480
 Db 421 TCCACTATTGGCACTACCCAGAAACAAGATTTAAAAAATAACAAAGTAACTAATCCACT 480
 Qy 481 CGAAGCTATCATGTATGTTTAAAGAAATCATATTTAAACCAACATCTCTTAAAAAA 540
 Db 481 CGAAGCTATCATGTATGTTTAAAGAAATCATATTTAAACCAACATCTCTTAAAAAA 540
 Qy 541 CAAGCATATTTCCAAAGAGACAATATGTATGATAGTTTAAACAATCTAAGCGACAAA 600
 Db 541 CAAGCATATTTCCAAAGAGACAATATGTATGATAGTTTAAACAATCTAAGCGACAAA 600
 Qy 601 TTAATTCGAAAGGTAAAGCTATGACGTTCAAGATTTTCTTTCAATCTTGTGTA 660
 Db 601 TTAATTCGAAAGGTAAAGCTATGACGTTCAAGATTTTCTTTCAATCTTGTGTA 660
 Qy 661 ATGTGTTTATATACATTTCTCTCTTCAATAGATGATTTCTTCCGATTTTAA 720
 Db 661 ATGTGTTTATATACATTTCTCTCTTCAATAGATGATTTCTTCCGATTTTAA 720
 Qy 721 ATGACTATAAAGTCATTTTATATAGAGACGACATGCTAGATTTCTCGTTCAAAAATC 780
 Db 721 ATGACTATAAAGTCATTTTATATAGAGACGACATGCTAGATTTCTCGTTCAAAAATC 780
 Qy 781 TTTCTGATTTTTTAAAGCTAGTTGGCAACCCGTTTCTTCAAGAAATTTGATTT 840
 Db 781 TTTCTGATTTTTTAAAGCTAGTTGGCAACCCGTTTCTTCAAGAAATTTGATTT 840
 Qy 841 TTCAAAAAAATAGTTTATTTCTCTTAAATAGAAACACTTAGAAAAATGAGT 900
 Db 841 TTCAAAAAAATAGTTTATTTCTCTTAAATAGAAACACTTAGAAAAATGAGT 900
 Qy 901 TGCAGACTAGCCCTAGAAATGTTTCCCAATTAATTAACAATCACTGATTAATTTTGG 960
 Db 901 TGCAGACTAGCCCTAGAAATGTTTCCCAATTAATTAACAATCACTGATTAATTTTGG 960
 Qy 961 GCCAGCCCCCAATTAATTTTAAACCGAAATGAGACGAAACCAATCTGAGCTAT 1020
 Db 961 GCCAGCCCCCAATTAATTTTAAACCGAAATGAGACGAAACCAATCTGAGCTAT 1020
 Qy 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAAATCAGTTTAAAGTCAATGTC 1080
 Db 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAAATCAGTTTAAAGTCAATGTC 1080
 Qy 1081 TGAAGATGCGGTTTGGCAACGATAGCCACCGTATCATAGTGTGCTCAAGTCA 1140
 Db 1081 TGAAGATGCGGTTTGGCAACGATAGCCACCGTATCATAGTGTGCTCAAGTCA 1140
 Qy 1141 GGTTCGGCAGCTCTCGTGTATCTCAATGCGATACATGCTTGTTCACACGTTCTGTC 1200
 Db 1141 GGTTCGGCAGCTCTCGTGTATCTCAATGCGATACATGCTTGTTCACACGTTCTGTC 1200
 Qy 1201 TTGTTCCATCGTCCAAAGCTTGGCTATTTGAAACCAAGAGATACGTAATCCCAACAT 1260
 Db 1201 TTGTTCCATCGTCCAAAGCTTGGCTATTTGAAACCAAGAGATACGTAATCCCAACAT 1260
 Qy 1261 CCATCTTACTCATGCACTTTCATGCAACACGCAATATGTTTCTGAAACCAATCCAT 1320
 Db 1261 CCATCTTACTCATGCACTTTCATGCAACACGCAATATGTTTCTGAAACCAATCCAT 1320
 Qy 1321 AAGATATCAACAAGCTAGCTTCTCCGCTAGCTTCTCTCTCTGCGATCTTTT 1380
 Db 1321 AAGATATCAACAAGCTAGCTTCTCCGCTAGCTTCTCTCTCTGCGATCTTTT 1380

Qy 1381 CGTCCACCA 1389
 Db 1381 CGTCCACCA 1389
 RESULT 4
 AX224395
 LOCUS AX224395 1394 bp DNA linear PAT 10-SEP-2001
 DEFINITION Sequence 2 from Patent WO0160997.
 ACCESSION AX224395
 VERSION AX224395.1 GI:15554637
 KEYWORDS
 SOURCE
 ORGANISM
 Zea mays
 Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 REFERENCE
 1
 AUTHORS Albertsen, M.C., Fox, T.W., Garnat, C.W., Huffman, G. and Kendall, T.L.
 TITLE Male tissue-preferred regulatory region and method of using same
 JOURNAL Patent: WO 0160997-A 2 23-AUG-2001;
 PIONEER HI-BRED INTERNATIONAL, INC. (US)
 FEATURES
 location/Qualifiers
 1..1394
 /organism="Zea mays"
 /mol_type="unassigned DNA"
 /db_xref="taxon:4577"
 ORIGIN
 Query Match 94.0%; Score 1311; DB 6; Length 1394;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 CGATGCTCTCTATGAAAAAGATGATCAATGCTATATCCGTTTCTTAAAGTCC 60
 Db 1 CGATGCTCTCTATGAAAAAGATGATCAATGCTATATCCGTTTCTTAAAGTCC 60
 Qy 61 CTTCCTTGGCCCTTATTAAGTGAATGCGGGTTCAAAAACTTCCAGGGGATGAT 120
 Db 61 CTTCCTTGGCCCTTATTAAGTGAATGCGGGTTCAAAAACTTCCAGGGGATGAT 120
 Qy 121 CTCGATGTTCACTTCTCCACCTCGCGTTGACATTTCTTGATGTCGATGTTCCCAT 180
 Db 121 CTCGATGTTCACTTCTCCACCTCGCGTTGACATTTCTTGATGTCGATGTTCCCAT 180
 Qy 181 CTGACCGAGGCCCATCAACACCTTTGGAGACCCATCAAGGGCTTTCCGATGGCCCA 240
 Db 181 CTGACCGAGGCCCATCAACACCTTTGGAGACCCATCAAGGGCTTTCCGATGGCCCA 240
 Qy 241 CGAAGCGTATGCGGTCGTGATGATCAGGGGATATATGTCGCCCAATCGTCACTATA 300
 Db 241 CGAAGCGTATGCGGTCGTGATATCAGGGGATATATGTCGCCCAATCGTCACTATA 300
 Qy 301 TTATTAATCTTTAGATTTTATTTTAAATTTTGGAAAAATTAACAACTTATCTTTGTGTA 360
 Db 301 TTATTAATCTTTAGATTTTATTTTAAATTTTGGAAAAATTAACAACTTATCTTTGTGTA 360
 Qy 361 GGGCTCAGCATGATTTTGGCTTAAAGGCGCCAGAAATGCGAGACCGCATGCTAGTG 420
 Db 361 GGGCTCAGCATGATTTTGGCTTAAAGGCGCCAGAAATGCGAGACCGCATGCTAGTG 420
 Qy 421 TCCACTATTGGCACTACCCAGAAACAAGATTTAAAAAATAACAAAGTAACTAATCCACT 480
 Db 421 TCCACTATTGGCACTACCCAGAAACAAGATTTAAAAAATAACAAAGTAACTAATCCACT 480
 Qy 481 CGAAGCTATCATGTATGTTTAAAGAAATCATATTTAAACCAACATCTCTTAAAAAA 540
 Db 481 CGAAGCTATCATGTATGTTTAAAGAAATCATATTTAAACCAACATCTCTTAAAAAA 540
 Qy 541 CAAGCATATTTCCAAAGAGACAATATGTATGATAGTTTAAACAATCTAAGCGACAAA 600
 Db 541 CAAGCATATTTCCAAAGAGACAATATGTATGATAGTTTAAACAATCTAAGCGACAAA 600

QY 601 TTAATCGAAGGTAAGCTATGACGTTCAATTTTCTTTTCAATCTGTATTTTGT 660
DB 601 TTAATCGAAGGTAAGCTATGACGTTCAATTTTCTTTTCAATCTGTATTTTGT 660
QY 661 ATTGTTTATATACATTTTCTTCTCTTACAAATAGATGATTTTCTCCGATTTTATATA 720
DB 661 ATTGTTTATATACATTTTCTTCTCTTACAAATAGATGATTTTCTCCGATTTTATATA 720
QY 721 ATGACTATAAGTCATTTTATATATAGAGACGATGTCGTAGATTCTGTTCAAAAATC 780
DB 721 ATGACTATAAGTCATTTTATATATAGAGACGATGTCGTAGATTCTGTTCAAAAATC 780
QY 781 TTTCTGATTTTATTAAGCTAGTTTGGCAACCCGTGTTCTTCAAAAGATTTTGAATTT 840
DB 781 TTTCTGATTTTATTAAGCTAGTTTGGCAACCCGTGTTCTTCAAAAGATTTTGAATTT 840
QY 841 TTCAAAAAAATTAAGTTATTTTCTCTTATTAATAATAGAAAAACATTAGAAAAATAGAT 900
DB 841 TTCAAAAAAATTAAGTTATTTTCTCTTATTAATAATAGAAAAACATTAGAAAAATAGAT 900
QY 901 TGCAGACTAGCCCTAGAAATGTTTTCCCAATTAATCAATCACTGTATTAATTTTGG 960
DB 901 TGCAGACTAGCCCTAGAAATGTTTTCCCAATTAATCAATCACTGTATTAATTTTGG 960
QY 961 GCCAGCCCCATTAATTTTAAACCGAACTGAATGAGACGAAACCAATCTGAGTAT 1020
DB 961 GCCAGCCCCATTAATTTTAAACCGAACTGAATGAGACGAAACCAATCTGAGTAT 1020
QY 1021 TTCTCTAGATTATTAATAAGGAGAGAGAGAGAAATCAGTTTAACTGATTTGCC 1080
DB 1021 TTCTCTAGATTATTAATAAGGAGAGAGAGAGAAATCAGTTTAACTGATTTGCC 1080
QY 1081 TGAAGTGTGGGTTTGGCAACGATGACCGTAATCAATGCTAATGAGTCTGATGCTCA 1140
DB 1081 TGAAGTGTGGGTTTGGCAACGATGACCGTAATCAATGCTAATGAGTCTGATGCTCA 1140
QY 1141 GGTTCGGACGCTCTCGGTGATCTCAATGAGCATCTGCTGTTTCAACCGTCTGTC 1200
DB 1141 GGTTCGGACGCTCTCGGTGATCTCAATGAGCATCTGCTGTTTCAACCGTCTGTC 1200
QY 1201 TTGTTCCATGCTCAAGCCTTGCTTATTTGAAACCAAGAGATACCTATCCCAACAT 1260
DB 1201 TTGTTCCATGCTCAAGCCTTGCTTATTTGAAACCAAGAGATACCTATCCCAACAT 1260
QY 1261 CCATCTTACTACTGCAACTTCATGCAACACGCACTATGTTCTGTAAC 1311
DB 1261 CCATCTTACTACTGCAACTTCATGCAACACGCACTATGTTCTGTAAC 1311

RESULT 5
BD062177 1394 bp DNA linear PAT 27-AUG-2002
LOCUS Male tissue-preferred regulatory region and method of using same.
DEFINITION BD062177.1 GI:22607782
ACCESSION BD062177.1
VERSION JP 2001520523-A/2.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1394)
Alberstein, M.C., Fox, T.W., Garneat, C.W., Huffman, G.A. and
Kendall, T.L.
Male tissue-preferred regulatory region and method of using same
PIONEER HI BRED INTERNATIONAL INC
JP JP 2001520523-A/2
COMMENT
BD 30-OCT-2001 JP 199504910
PR 19-JUN-1998 JP 08/880499
PI MARC C ALBERTSEN, TIMOTHY W FOX, CARL W GARNAAT, GARY A HUFFMAN,
TIMMY L KENDALL

PC C12N15/82, C12N15/29, C12N9/24, C12N9/22, C12N9/10, C12N9/00 PC
C07K14/34, C1201/68,
PC A01H5/00
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers.
FEATURES
source
1..1394
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 94.0%; Score 1311; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCATGGTGTCTATATAGAAAAAGATGATCAATGTGCTATATCCGTTTCTTAGGGTCC 60
DB 1 CCATGGTGTCTATATAGAAAAAGATGATCAATGTGCTATATCCGTTTCTTAGGGTCC 60
QY 61 CTTCTTCTGCTTATTAAGTGAATCGGGTTACAAAAAATTTCACGGGTGATGAT 120
DB 61 CTTCTTCTGCTTATTAAGTGAATCGGGTTACAAAAAATTTCACGGGTGATGAT 120
QY 121 CTCGATGTTCCACTTCTCCCACTCGCGTTGCACTTTCTTGGAATGCGGTGTTCCAT 180
DB 121 CTCGATGTTCCACTTCTCCCACTCGCGTTGCACTTTCTTGGAATGCGGTGTTCCAT 180
QY 181 CTGACCGAGGCCCATGACGACCTTTGCGGACACCCATCAAGGACCTTTGCGATGGCCCA 240
DB 181 CTGACCGAGGCCCATGACGACCTTTGCGGACACCCATCAAGGACCTTTGCGATGGCCCA 240
QY 241 CGAGAGATATCGGGTGTGATATCCAGGGGATATATGTCCCGACAAATGTCACCTATA 300
DB 241 CGAGAGATATCGGGTGTGATATCCAGGGGATATATGTCCCGACAAATGTCACCTATA 300
QY 301 TTAATATCTTATATATATTTTATTTTGGAAAAATTAACAACCTTATATCTTTGTGTA 360
DB 301 TTAATATCTTATATATATTTTATTTTGGAAAAATTAACAACCTTATATCTTTGTGTA 360
QY 361 GGGCCTCAGATGATTTTTCGCTTAGGGCCAGAAATGCGAGGACAGCCATGCTTAGTG 420
DB 361 GGGCCTCAGATGATTTTTCGCTTAGGGCCAGAAATGCGAGGACAGCCATGCTTAGTG 420
QY 421 TCCACTATTGGCACTACCCAGAACAGATTTAAAAAATTAACCAAGTAACTAATCACT 480
DB 421 TCCACTATTGGCACTACCCAGAACAGATTTAAAAAATTAACCAAGTAACTAATCACT 480
QY 481 CGAAGCTATCATGTATATGTTTAAAGAAACATCTATTTAAACGACATCTCTTAAAAAA 540
DB 481 CGAAGCTATCATGTATATGTTTAAAGAAACATCTATTTAAACGACATCTCTTAAAAAA 540
QY 541 CAAGCATATTTGCAAGAGACAAATATGTTTCAAGTTTCAAAACATCTAAGAGCGACAA 600
DB 541 CAAGCATATTTGCAAGAGACAAATATGTTTCAAGTTTCAAAACATCTAAGAGCGACAA 600
QY 601 TTAATCGAAGGTAAGCTATGACGTTCAATTTTCTTTTCAATCTGTATTTTGT 660
DB 601 TTAATCGAAGGTAAGCTATGACGTTCAATTTTCTTTTCAATCTGTATTTTGT 660
QY 661 ATTGTTTATATACATTTTCTTCTCTTACAAATAGATGATTTTCTCCGATTTTATATA 720
DB 661 ATTGTTTATATACATTTTCTTCTCTTACAAATAGATGATTTTCTCCGATTTTATATA 720
QY 721 ATGACTATAAGTCATTTTATATATAGAGACGATGTCGTAGATTCTGTTCAAAAATC 780
DB 721 ATGACTATAAGTCATTTTATATATAGAGACGATGTCGTAGATTCTGTTCAAAAATC 780
QY 781 TTTCTGATTTTATTAAGCTAGTTTGGCAACCCGTGTTCTTCAAAAGATTTTGAATTT 840
DB 781 TTTCTGATTTTATTAAGCTAGTTTGGCAACCCGTGTTCTTCAAAAGATTTTGAATTT 840

RESULT 6				
AX224402				
LOCUS				
DEFINITION	AX224402	255 bp	DNA	linear
ACCESSION	Sequence 9	from Patent WO0160997.		
AX224402				
VERSION				
KEYWORDS	AX224402.1	GI:15554644		
SOURCE				
ORGANISM				
	Zea mays			
	Zea mays			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD			
	clade; Panicoideae; Andropogoneae; Zea.			
REFERENCE				
AUTHORS	1			
TITLE	Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.			
JOURNAL	Male tissue-preferred regulatory region and method of using same			
	Patent: WO 0160997-A 9 23-AUG-2001;			
	PIONEER HI-BRED INTERNATIONAL, INC. (US)			
FEATURES				
source	location/Qualifiers			
	1..255			
	/organism="Zea mays"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:4577"			

RESULT 7				
AX224396				
LOCUS				
DEFINITION	158 bp	DNA	linear	PAT 10-SEP-2001
ACCESSION	Sequence 3	from Patent WO0160997.		
AX224396				
VERSION	AX224396.1	GI:15554638		
KEYWORDS				
SOURCE				
ORGANISM				
	Zea mays			
	Zea mays			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
	Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD			
	clade; Panicoideae; Andropogonaceae; Zea.			
REFERENCE	1			
AUTHORS	Albertsen, M.C., Fox, T.W., Garnaat, C.W., Huffman, G. and Kendall, T.L.			
TITLE	Male tissue-preferred regulatory region and method of using same			
JOURNAL	Patent: WO 0160997-A 3 23-AUG-2001;			
	PIONEER HI-BRED INTERNATIONAL, INC. (US)			
FEATURES				
source	1..158			
	Location/Qualifiers			
	/organism="Zea mays"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:4577"			

LOCUS	AX224398	50 bp	DNA	linear	PAT 10-SEP-2001
DEFINITION	Sequence 5 from Patent WO0160997.				
ACCESSION	AX224398				
VERSION	AX224398.1	GI:15554640			
KEYWORDS					
SOURCE	Zea mays				
ORGANISM	Zea mays				
	Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.				
REFERENCE	1				
AUTHORS	Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.				
TITLE	Male tissue-preferred regulatory region and method of using same				
JOURNAL	Patent: WO 0160997-A 5 23-AUG-2001;				
FEATURES	PIONEER HI-BRED INTERNATIONAL, INC. (US)				
	location/Qualifiers				

```

ORIGIN
Query Match 3.6%; Score 50; DB 6; Length 50;
Best Local Similarity 100.0%; Pred. No. 1,8e-14;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1229 CTGAACCAAGAGTACCTTACTCCAAACAATCATCTTACTATGCAC 1278
      |||||
Db 1 CTGAACCAAGAGTACCTTACTCCAAACAATCATCTTACTATGCAC 50

```

```

LOCUS       AX224399             40 bp      DNA      linear      PAT 10-SEP-2001
DEFINITION   Sequence 6 from Patent WO0160997.
ACCESSION    AX224399.1      GI:15554641
KEYWORDS     Zea mays
SOURCE       Zea mays
ORGANISM     Zea mays

REFERENCE    1
AUTHORS      Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
TITLE        Male tissue-preferred regulatory region and method of using same
JOURNAL      PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES     Location/Qualifiers
             1..40
             /organism="Zea mays"
             /mol_type="unassigned DNA"
             /db_xref="taxon:4577"

ORIGIN
Query Match      2.3%; Score 40; DB 6; Length 40;
Best Local Similarity 100.0%; Pred. No. 3.4e-09;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY
1239 AGGATACCTACTGCTCCAAACAATCTTACTGATGCAAC 1278
|||||
1 AGGATACCTACTGCTCCAAACAATCTTACTGATGCAAC 40

Db
1239 AGGATACCTACTGCTCCAAACAATCTTACTGATGCAAC 1278
|||||
1 AGGATACCTACTGCTCCAAACAATCTTACTGATGCAAC 40

RESULT 10
LOCUS       AX224397             30 bp      DNA      linear      PAT 10-SEP-2001
DEFINITION   Sequence 4 from Patent WO0160997.
ACCESSION    AX224397
KEYWORDS     Zea mays
SOURCE       Zea mays
ORGANISM     Zea mays

REFERENCE    1
AUTHORS      Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
TITLE        Male tissue-preferred regulatory region and method of using same
JOURNAL      PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES     Location/Qualifiers
             1..30
             /organism="Zea mays"
             /mol_type="unassigned DNA"
             /db_xref="taxon:4577"

ORIGIN
Query Match      2.3%; Score 30; DB 6; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.00067;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY
1179 CATGCTTTTCACCGTCTGCTTTGTTCCA 1208
|||||
1 CATGCTTTTCACCGTCTGCTTTGTTCCA 30

Db
1 CATGCTTTTCACCGTCTGCTTTGTTCCA 30

RESULT 11
LOCUS       AC149836             114817 bp      DNA      linear      HTG 22-JUN-2004
DEFINITION   Zea mays clone ZMMB0496L17, *** SEQUENCING IN PROGRESS ***, 3
ACCESSION    AC149836
VERSION      AC149836.1      GI:49035067
KEYWORDS     HTG; HTGS_PHASE1; HTGS_FULFLOP; HTGS_ACTIVEFIN.
SOURCE       Zea mays

```

```

ORGANISM     Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE    1 (bases 1 to 114817)
AUTHORS      Birren,B., Nussbaum,C., Lander,E., Butler,E., Wang,R., Bharti,A.K.
and Messing,J.
TITLE        Zea mays, clone ZMMB0496L17
JOURNAL      Unpublished
AUTHORS      2 (bases 1 to 114817)
             Birren,B., Nussbaum,C., Lander,E., Butler,E., Wang,R., Bharti,A.K.,
             Messing,J., Abouelleil,A., Allen,N., Anderson,M., Anderson,S.,
             Arachchi,H.M., Barua,N., Baetjen,V., Bloom,T., Boguslavsky,L.,
             Borkgatter,B., Camarata,J., Chang,J., Choepel,Y., Collimore,A.,
             Cook,A., Cooke,P., Corum,B., Dearlilano,K., Diaz,J.S., Dodge,S.,
             Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P.,
             Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L.,
             Grand-pierre,N., Hafez,N., Hagopian,D., Hages,B., Hall,J.,
             Horvath,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
             Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
             Liu,G., Liu,X., Lui,A., Mabbitt,R., Maclean,C., Macdonald,P.,
             Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrum,J.,
             Meneses,L., Minova,T., Miengo,V., Murphy,T., Naylor,J., Nguyen,C.,
             Nguyen,T., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
             O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
             Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
             Roman,J., Schauer,S., Schupbach,R., Seaman,S., Severy,P., Smith,C.,
             Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
             Talmes,J., Tsefaye,S., Theodore,J., Topham,K., Travers,M.,
             Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
             Wyman,D., Young,G., Zainoun,J., Zemke,L., Zimmer,A. and Zody,M.
             Direct Submission
             Submitted (22-JUN-2004) Whitehead Institute/MIT Center for Genome
             Research, 320 Charles Street, Cambridge, MA 02141, USA
             All repeats were identified using RepeatMasker:
             Smit,A.F.A. & Green,P. (1996-1997)
             http://ftp.genome.washington.edu/RM/RepeatMasker.html
             ----- Genome Center
             Center: Whitehead Institute/MIT Center for Genome Research
             Center code: MIBR
             Web site: http://www-seq.wi.mit.edu
             Contact: sequence_submissions@broad.mit.edu
             -----
             Bharti, AK and Messing, J: The Plant Genome Initiative at
             Rutgers, Waksman Institute, Rutgers, The State University of New
             Jersey, 190 Frelinghuysen Road, Piscataway, NJ 08854, USA
             (http://pgr.rutgers.edu)
             Butler, E and Wang, R: Arizona Genomics Institute, Biological
             Sciences West, 448A, P.O. Box 210088, University of Arizona,
             Tucson, AZ 85721, USA (http://www.genome.arizona.edu)
             ----- Project Information
             Center project name: L30352
             Center clone name: 496_L_17
             -----
             * NOTE: This is a 'working draft' sequence. It currently
             * consists of 3 contigs. The true order of the pieces
             * is not known and their order in this sequence record is
             * arbitrary. Gaps between the contigs are represented as
             * runs of N, but the exact sizes of the gaps are unknown.
             * This record will be updated with the finished sequence
             * as soon as it is available and the accession number will
             * be preserved.
             *
             1       35156: contig of 35156 bp in length
             *       35157       35256: gap of 100 bp
             *       35257       97940: contig of 62684 bp in length
             *       97941       98040: gap of 100 bp
             *       98041      114817: contig of 16777 bp in length.
             *
             Location/Qualifiers
             1..114817
             /organism="Zea mays"
             /mol_type="genomic DNA"
             /culivar="B73"
             /db_xref="taxon:4577"

```

ORIGIN /clone="ZMMBC0496L17" /clone_lib="CHORI-201 Maize B73 BAC Mbol Library"

Query Match 1.8%; Score 25; DB 2; Length 114817;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 844 AAAAAAAAAATGATTATTTCTCTT 868
Db 27886 AAAAAAAAAATGATTATTTCTCTT 27910

RESULT 12
AE010607
LOCUS
DEFINITION AE010607 10384 bp DNA linear BCT 25-MAR-2002
Fusobacterium nucleatum subsp. nucleatum ATCC 25586, section 149 of
197 of the complete genome.
ACCESSION AE010607 AE009951
VERSION AE010607.1 GI:19714578
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
SOURCE

1. (bases 1 to 10384)
Kaprali,V., Anderson,I., Ivanova,N., Reznik,G., Los,T.,
Lykidis,A., Bhattacharyya,A., Bartman,A., Gardner,W., Grechkin,G.,
Zhu,L., Vasileva,O., Chu,L., Kogan,Y., Chaga,O., Goldsman,E.,
Bernal,A., Larsen,N., D'Souza,M., Walunas,T., Pusch,G.,
Hasselkorn,R., Fontein,M., Kyriades,N. and Overbeek,R.
Genome sequence and analysis of the oral bacterium Fusobacterium
nucleatum strain ATCC 25586
J. Bacteriol. 184 (7), 2005-2018 (2002)
2186394
11889109
2 (bases 1 to 10384)
Kaprali,V., Anderson,I., Ivanova,N., Reznik,G., Los,T.,
Lykidis,A., Bhattacharyya,A., Bartman,A., Gardner,W., Grechkin,G.,
Zhu,L., Chu,L., Kogan,Y., Chaga,O., Goldsman,E., Bernal,A.,
Larsen,N., D'Souza,M., Walunas,T., Pusch,G.D., Hasselkorn,R.,
Fontein,M., Kyriades,N. and Overbeek,R.
Direct Submission
Submitted (13-FEB-2002) Integrated Genomics, 2201 W. Campbell Park
Drive, Chicago, IL 60612, USA
Location/Qualifiers
1. .10384
/organism="Fusobacterium nucleatum subsp. nucleatum ATCC
25586"
/mol_type="genomic DNA"
/strain="ATCC 25586"
/sub_species="nucleatum"
/db_xref="ATCC:25586"
/db_xref="taxon:190304"
162. .1664
/gene="FN0998"
162. .1664
/gene="FN0998"
/codon_start=1
/transl_table=11
/product="Dipeptide-binding protein"
/protein_id="AAL95194.1"
/db_xref="GI:19714579"
/translation="MKKIFYLFLVSLFLVGGCGKNTSPNGSTVVIIGGAKPKSIDP
HMYNSIPDLVSRQFYNTLPFSREKDSIRPELASEYRKYNDKLDIILKKGKPHDGS
ELTADVVFSFERMKDPGSSIVVESIDVEKVDYELFLKSSSALLVNLAPIT
SIVNKRYEAGNDLSVAPMGTAFLIAYNDGEKLELAEKQYFEGAPVEKEITPFI
PEDTSMALLETGEVDIATGMPVSTQITTEANDKLDLISEPTATETVCLINVEKTFPI
NKDFRKLNAIADKOSIIDISIFSGRGVAKSIVPNVFGYDGFEEYFPNPEAKELI
EKSQVDMSPSLVYNDSPVRLVAQIIONLKIDVGIIDMIETLEMGTYLOKTEGDFS
AYLGWISGTSDDIVLYPLDLSKSGIFGNRRARYSNPREDKEVEKARIALKPEERKE
HFNNAQIIISOEDPLIVLYNKNENIGINKRIKGEYDPTTMKFKNLEIK"

CDS
/gene="FN0999"
1669. .2712
/gene="FN0999"
/EC_number="3.4.11.-"
/codon_start=1
/transl_table=11
/product="Deblocking aminopeptidase"
/protein_id="AAL95195.1"
/db_xref="GI:19714580"
/translation="MILMKNDIKYTLAKTVELINIPSPVGTTHNAIEVYQELKGLGI
KNYITTKGALIAVTKGDSYKKMI SHAVDLGVAVKRIKKSARLEVTNNGGFAMGS
VEGENVTIHTISGTYSGTLPLIKASVYVDVAEMPRTERTEIRIDEDKTEDV
LKGLIQCDFISFETHTRIIDNGYIKSRYLDDKLCVAQILSYIKLKNNKLPKPDLY
IYFSNYEIGHGVSPEDLDEFLIYVDIGLVAGEDAHDGKVOJIIAKOSRETYFTL
RKLOETASKNNIKTYGVNRYGSDATTAVALGDFPYACIGPVNDAHTHERCHND
GIETVKKLIAYL"
2854. .3936
/gene="FN1000"
2854. .3936
/gene="FN1000"
/EC_number="2.8.1.6"
/codon_start=1
/transl_table=11
/product="Biotin synthase"
/protein_id="AAL95196.1"
/db_xref="GI:19714581"
/translation="MLKENSAGCGKFNFPFLKEKNNNOAEPNSVKEFISYLNKII
NEKYITREBAIFLSQIPNNMETNIIIPNADOIRBVPCCGYPLCTIINAKSKCS
ENCKTCAOSVHKGTIDVGLISKELAEAKRNENBAHRPSLTSGRLGNKEKL
DKLVIYRIYIEBHSKELCASHGICTEALQKVDACVALYHNHLESSRRYPVCT
SHSYDRIINITKNAVGLDVCSGGIFGLGETIEDRIDMALDLRALEHSYDINVLPT
IPGTFPENNEVNPEILKTIISYIFIMPKFLRYCGGRIRKGEHAKTGLRCGINSAL
TGNFLTGTITETDKNMKEIGVEI"
3926. .4585
/gene="FN1001"
3926. .4585
/gene="FN1001"
/EC_number="6.3.3.3"
/codon_start=1
/transl_table=11
/product="Dechobiotin synthetase"
/protein_id="AAL95197.1"
/db_xref="GI:19714582"
/translation="MKFDPFVIGTDTGVKTYVSTLLYKALKRYNFOYKXIOSGCF
LKDGKLTAPDVDFLTKFVGIYDOSMTYTLKEVSPHASEBGEIRIEIENVKEHYE
DLKKGHSNIVBAGAGLYVPLIRDI FYIYDIKLFNLVVIIVCCGKVAINTMTLTN
ALDTWGIKDHGLVFNRYRGQFFEDNIRVALLSKIKNYLIITKNGQKEISDEKTEKF
N"
4653. .6011
/gene="FN1002"
4653. .6011
/gene="FN1002"
/EC_number="2.6.1.62"
/codon_start=1
/transl_table=11
/product="Adenosylmethionine-8- amino-7-oxononanoate
aminotransferase"
/protein_id="AAL95198.1"
/db_xref="GI:19714583"
/translation="WRGVCTIMINNLSELOKDKLVYFHPCAQMKDPEENPPLVYKKGD
GLYLIDENGNKYMDCISSWVNLFGCHCKRINRIITTEOVNLI,EHVIFANFTHEPAEL
CEELTKVLPJGINKFLPSDNGSSCIEMALKSPYHLOTGNPOKTFISLENAVHGET
IGALGVGDVDFIETERYPLIEGRKRVYIDSKLNDPVLDECKLELDLIIKN
HNETACMVVEPMVOCAAKIKIYSAKELAAARDITKXNIIH,IIIDSEIAMGFGTGCMFA
CEHAIETEDMACIKAGSSGYPIIAMLCTITDIFNAFADYDKRSFLASHITYSNPL
GCRNLAEVLRIPEKRNIIKTINEKGTYNANMKELFEEDKSYIKQIRNIGL,GAILEKD
KILPVRIGKEIYFALKKGAFLRPIGNSVYFMPYVITTEIDMKMLHVESIBEL
N"
6737. .7558
/gene="FN1003"
6737. .7558
/gene="FN1003"
/codon_start=1


```

* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 127823: contig of 127823 bp in length
* 127824 127923: gap of unknown length
* 127924 147441: contig of 19518 bp in length
* 147442 147541: gap of unknown length
* 147542 148509: contig of 36968 bp in length.
FEATURES
    source
        1..184509
            /organism="Canis familiaris"
            /mol_type="genomic DNA"
            /db_xref="taxon:9615"
            /clone="RP81-70312"
            /clone_1db="RP81"
            1..127823
            /note="assembly_fragment"
            clone_end:T7
            vector_side:left"
        misc_feature
            1..3358
            /note="clone overlaps with GenBank Accession Number
            AC110671 clone RP81-69D24 (center project name aw)"
            127924..147441
            /note="assembly_fragment"
            129971..184509
            /note="clone overlaps with GenBank Accession Number
            AC13567 clone RP81-265M12 (center project name cio)"
            147542..184509
            /note="assembly_fragment"
            clone_end:SP6
            vector_side:right"
ORIGIN
Query Match 1.7%; Score 24; DB 2; Length 184509;
Best Local Similarity 100.0%; Pred.No. 1;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1036 AAAAGGAGAGAGAGAGAGAGAA 1059
DB 23292 AAAAGGAGAGAGAGAGAGAA 23269
RESULT 14
AC098210/c 212621 bp DNA linear HTG 13-MAY-2003
LOCUS AC098210
DEFINITION Rattus norvegicus clone CH230-176E2, WORKING DRAFT SEQUENCE.
ACCESSION AC098210.8 GI:30581396
VERSION HTG; HTGS PHASE2; HTGS DRAFT; HTGS_FULLTOP.
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
            1 (bases 1 to 212621)
            Muzny,D.,Marle,, Metzker,M.,Lee,, Abramson,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Aisbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,I.M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Bismail,K., Blair,J., Blankenburg,K., Blych,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavaros,I., Caesar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dedrich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Diya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Bayes,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Frazer,C.M., Gabler,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,

```

```

Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Haylak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hognes,M.,
Hollins,B., Howells,S., Hulyk,S., Hune,V., Idlebit,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolyet,A.,
Karcachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,D.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorenshew,L., Loulseged,H., Lozdo,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindaratne,M., Mahmood,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,B.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Mundiass,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Parks,K.,
Nwackemehe,O., Okwodu,G., Olarnunagsoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfankuch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.,
Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,M., Savery,G., Scherer,S., Scott,G., Shatman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sison,I., Sitter,C.D., Smejls,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steinle,M., Strong,R., Sutton,A., Svatek,A., Tabox,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Umami,K.,
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wiczcyk,R., Wooden,H., Wotley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhausen,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 212621)
Worley,K.C.
Direct Submission
Submitted (13-OCT-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 212621)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 13, 2003 this sequence version replaced gi:23288092.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
----- Project Information
Contact: hgsc@bcm.tmc.edu
Center project name: GILN
Center clone name: CH230-176E2
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 205647 bases at least Q40
Consensus quality: 207827 bases at least Q30
Consensus quality: 209298 bases at least Q20

```


	Query Match	1.7%; Score 24; DB 2; Length 212621;
	Best Local Similarity	100.0%; Pred. No. 1;
	Matches	24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy	848 AAATAGTTATTCTTCTTTAT 871	
Dd	1464 AAAATGATTATTCTTCTTTAT 1441	
RESULT 15		
AC110715/c		
LOCUS	AC110715	213753 bp DNA linear HTG 19-NOV-2002
DEFINITION	Rattus norvegicus clone CH230-155J23. *** SEQUENCING IN PROGRESS	
KEYWORDS	***, 2 unordered pieces.	
SOURCE	AC110715 GI:25072769	
ORGANISM	HTG_PHASE1; HTGS_DRAFT; HTGS_ENRICHED. Rattus norvegicus (Norway rat) Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	
REFERENCE	1 (bases 1 to 213753)	
AUTHORS	Munuy,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J., Allen,C., Allen,H., Alb Brooks,S., Amin,A., Angiano,D., Aryalabechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaramaika,D., Barber,M., Barneshead,M., Benahmed,F., Biwaolo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Byrant,N., Buhan,C., Burch,P., Butrell,K., Calderon,E., Cadenas,V., Carter,K., Cavazos,I., Caesar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., P.Souza,L., Daviila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denison,S., Derxmo,C., Ding Y., Dinh,H., Divya,K., Diexper,H., Dugan-Rocha,S., Dunn,A., Durdin,K., Duval,B., Eaves,K., Egan,A., Bacetto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Flinley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Freder.C.M., Gabisi,A., Gante,R., Garcia,A., Garner,T., Garza,M., Georgievgeorgis,E., Geer,K., Gill,R., Grady,W., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamli,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,J., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins.B., Howells,S., Huliyk,S., Hume,J., Idlehard.D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson.R., Jolicet,A., Katchepy,S., Kelly,S., Khan,Z., King,L., Kovat,C.,	

Kwiatkowski, C., Kraft, C.L., Lebowitz, H., Lewan, J., Lewis, L., Li, Z., Liu, J.,
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
 Lorenzen, L., Louie, L., Louie, L., Louie, L., Louie, L., Louie, L.,
 Maitland, M., Mahindaratne, M., Mahmood, M., Mallory, K., Mangum, A.,
 Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
 Mahoney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
 Miossajevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 Nwokile, O., Okunolu, G., Olariu, A., Pal, S., Parks, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
 Plopper, F., Polidexter, A., Popovic, D., Primus, E., Pu, L.,
 Puzo, M., Quirio, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richard, S., Riggs, F.,
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.,
 Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smalls, D.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Soza, J.,
 Steinle, M., Strong, R., Sutton, A., Swack, A., Taber, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, E., Usmani, K.,
 Valdes, R., Vera, V., Villaseca, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wlezyk, R., Woodson, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhansen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
 Weinstein, G., and Gibbs, R.A.
 Unpublished
 Direct Submission
 2 (bases 1 to 213753)
 Morley, K.C.
 Direct Submission
 Submitted (15-FEB-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 213753)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (19-NOV-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Nov 19, 2002 this sequence version replaced gi:22689201.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GSTM
 Center clone name: CH230-15J23
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 203043 bases at least Q40
 Consensus quality: 204685 bases at least Q30
 Consensus quality: 205672 bases at least Q20
 Estimated insert size: 206221; sum-of-contigs estimation
 Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently

* consists of 2 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 212625: contig of 212625 bp in length
 * 212626 212725: gap of unknown length
 * 212726 213753: contig of 1028 bp in length.
 Location/Qualifier

FEATURES
 source

1. .213753
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-15J23"
 1. .1535
 /note="wgs end_extension
 clone_end:Sp6"
 1586. .2703
 /note="wgs end_extension
 clone_end:Sp6"
 5028. .5742
 /note="clone boundary
 clone_end:Sp6
 site:
 end_sequence:BH364626"
 33891. .35390
 /note="wgs_contig"
 211446. .212625
 /note="wgs_contig"

ORIGIN

Query Match 1.7%; Score 24; DB 2; Length 213753;
 Best Local Similarity 100.0%; Pred.No. 1;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 848 AAAATGATTATTTCTTTAT 871
 |||||
 Db 65390 AAAATGATTATTTCTTTAT 65367

Search completed: September 16, 2005, 15:25:37
 Job time : 4110 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2005, 08:08:50 ; Search time 578.5 Seconds
(without alignments)
14264.684 Million cell updates/sec

Title: US-10-713-381-1

Perfect score: 1394
Sequence: 1 ccatggtgtctctcatgaataa.....cttttcgtccaccaccatg 1394

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

1: N_Geneseq_16Dec04:*
2: genebegn1980s:*
3: genebegn1980s:*
4: genebegn2000s:*
5: genebegn2001as:*
6: genebegn2001bs:*
7: genebegn2002as:*
8: genebegn2002bs:*
9: genebegn2003as:*
10: genebegn2003bs:*
11: genebegn2003ds:*
12: genebegn2004as:*
13: genebegn2004bs:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1394	100.0	1394	2	AAK07408
2	1394	100.0	1394	5	AAH76332
3	1311	94.0	1394	2	AAK07409
4	1311	94.0	1394	5	AAH76333
5	157	11.3	158	5	AAH76340
6	111	8.0	158	5	AAH76334
7	50	3.6	50	5	AAH76336
8	40	2.9	40	5	AAH76337
9	30	2.2	30	5	AAH76335
10	23	1.6	849	5	ABL40437
11	21	1.5	333	5	ABV56505
12	21	1.5	459	4	ABAS8867
13	21	1.5	459	4	AAI38581
14	21	1.5	459	4	AAK3768
15	21	1.5	459	4	AAK07027
16	21	1.5	459	4	ABK32487
17	21	1.5	459	4	ABK07027
18	21	1.5	965	6	ABK07027
19	21	1.5	9255	6	ABK34152
20	21	1.5	9265	6	AAK63354

21	21	1.5	18817	6	ABK34494	ABK34494 Human met
22	21	1.5	18817	6	ABK70161	ABK70161 Chemically
23	21	1.5	18817	7	ADK97975	ADK97975 Bisluphit
24	21	1.5	26493	12	ADJ12386	ADJ12386 DNA fragm
25	20	1.4	260	5	ABV60726	ABV60726 Human pro
26	20	1.4	680	6	ABO18134	ABO18134 Oligonuc
27	20	1.4	680	6	ABO18135	ABO18135 Oligonuc
28	20	1.4	2523	2	AAK76376	AAK76376 Zea maye
29	20	1.4	10439	6	AAK47817	AAK47817 Alternati
30	20	1.4	11779	6	ACN42891	ACN42891 Maize sug
31	20	1.4	51198	11	ACN45000	ACN45000 Mouse gen
32	20	1.4	91071	11	ACN44004	ACN44004 Mouse gen
33	20	1.4	110000	6	ABA90193_2	Continuation (3 of
34	20	1.4	110000	6	ABA90193_3	Continuation (4 of
35	20	1.4	110000	6	ABQ87681_2	Continuation (3 of
36	20	1.4	110000	6	ABQ87681_3	Continuation (4 of
37	20	1.4	110000	8	ABK33717_2	Continuation (3 of
38	20	1.4	110000	8	ABK33717_3	Continuation (4 of
39	20	1.4	148497	12	ADQ18925	ADQ18925 Human scf
40	20	1.4	194534	12	ADQ97481	ADQ97481 Human can
41	20	1.4	337344	13	ABD32715	ABD32715 Human can
42	19	1.4	418	13	ADR65258	ADR65258 Cotton CD
43	19	1.4	468	2	AAK19065	AAK19065 Human PPA
44	19	1.4	485	4	ABA57528	ABA57528 Human foe
45	19	1.4	485	4	AAI37083	AAI37083 Probe #57

ALIGNMENTS

RESULT 1	AAK07408	standard; DNA; 1394 BP.
ID	AAK07408	
AC	AAK07408;	
XX	XX	
DT	08-JUN-1999	(first entry)
XX	XX	
DE	Zea maye M45 male tissue-preferred regulatory region.	
XX	XX	
XX	M45; male; tissue-preferred; regulatory region; plant cells;	
KW	plant tissue; differentiated; maize; hybrid seed; fertility; ss.	
OS	Zea maye.	
XX	XX	
PN	WO9859061-A1.	
XX	XX	
PD	30-DEC-1998.	
XX	XX	
PP	19-JUN-1998;	98WO-US012895.
XX	XX	
PR	23-JUN-1997;	97US-00880499.
XX	XX	
PA	(PION-) PIONEER HI-BRED INT INC.	
PI	Albertsen MC, Fox TW, Garnaat CW, Huffman GA, Kendall TL;	
XX	XX	
DR	WPI; 1999-105628/09.	
XX	XX	
PT	New nucleic acid encoding a M45 male tissue-preferred regulatory region	
XX	XX	
XX	- useful in mediating plant fertility, especially hybrid seed production.	
XX	XX	
XX	Claim 2; Page 22-23; 39pp; English.	
XX	XX	
CC	The sequence is that encoding an M45 male tissue-preferred regulatory	
CC	region. It may be used in the construction of a vector for a method of	
CC	producing exogenous genes in a male tissue-preferred manner, which is	
CC	useful in restoring or conferring fertility, such as in hybrid seed	
CC	production. In conferring fertility, a monocot/dicot plant is transformed	
CC	with the exogenous nucleotide sequence (a male sterility gene, preferably	
CC	M45), which encodes a product selected from auxins, rols and diptheria	
CC	toxins. Hybrid seeds are produced by cross-pollinating maize male fertile	
CC	and infertile plants	

```

XX Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;
SQ
Query Match      100.0%; Score 1394; DB 2; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCATGGGTCTCTATGAAAAAGATGAGTACATGTGTCTATATCCGTTTTCTTAAGGGTCC 60
Db 1 CCATGGGTCTCTATGAAAAAGATGAGTACATGTGTCTATATCCGTTTTCTTAAGGGTCC 60

Qy 61 CTTCCTTCGCTTATTAATCTGATGAAATCGGGGTTTACAAAAAATTCCACGGGTGCAATGAT 120
Db 61 CTTCCTTCGCTTATTAATCTGATGAAATCGGGGTTTACAAAAAATTCCACGGGTGCAATGAT 120

Qy 121 CTGCAATGTTCCACTCTCCCACTCGCGGTGCACTTTCTTGAAATGTCGGTGGTCCCAT 180
Db 121 CTGCAATGTTCCACTCTCCCACTCGCGGTGCACTTTCTTGAAATGTCGGTGGTCCCAT 180

Qy 181 CTGACCGAGGCCCATCAGACACCTTTGGGACACCCATCAAGGGCCCTTTGGAATGGCCCA 240
Db 181 CTGACCGAGGCCCATCAGACACCTTTGGGACACCCATCAAGGGCCCTTTGGAATGGCCCA 240

Qy 241 CGAAGCGTATCGGGTCGTGTGATCGAGGGATATATGTCCCCACAAATGTCACCTTATA 300
Db 241 CGAAGCGTATCGGGTCGTGTGATCGAGGGATATATGTCCCCACAAATGTCACCTTATA 300

Qy 301 TTATATATCTTTAGATTTTAAATTTTGAAGAAAAATACAAACTTATATCTTTGTGTA 360
Db 301 TTATATATCTTTAGATTTTAAATTTTGAAGAAAAATACAAACTTATATCTTTGTGTA 360

Qy 361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCAGAAATGCGAGACACGACCATGTCTAGTG 420
Db 361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCAGAAATGCGAGACACGACCATGTCTAGTG 420

Qy 421 TCCACTATTGGCACTACCCAGAACAGATTTAAAAAAATACAAAGTAACTATCCACT 480
Db 421 TCCACTATTGGCACTACCCAGAACAGATTTAAAAAAATACAAAGTAACTATCCACT 480

Qy 481 CGAAAGCTATCATGTATATGTTTAAAGAAACATCATTTAAACCACTCTCTTAAAAAA 540
Db 481 CGAAAGCTATCATGTATATGTTTAAAGAAACATCATTTAAACCACTCTCTTAAAAAA 540

Qy 541 CAAGCATATTTGCAAAAGACAAATTAATGTTACAGTTTACAAACATCTAAGACGACAAA 600
Db 541 CAAGCATATTTGCAAAAGACAAATTAATGTTACAGTTTACAAACATCTAAGACGACAAA 600

Qy 601 TTATATGAAAAGTAAAGCTATGACGTTCAAGATTTTCTTTTCAATCTGTATATTTGTT 660
Db 601 TTATATGAAAAGTAAAGCTATGACGTTCAAGATTTTCTTTTCAATCTGTATATTTGTT 660

Qy 661 ATTGTTTTTATATACATTTCTTCTTACATATGAGTATTTCTTCCGATTTTAA 720
Db 661 ATTGTTTTTATATACATTTCTTCTTACATATGAGTATTTCTTCCGATTTTAA 720

Qy 721 ATGACTATAAAGTCATTTTATATAAGAGCAGCATGTCGTAGATTTCTCGTCAAAAAATC 780
Db 721 ATGACTATAAAGTCATTTTATATAAGAGCAGCATGTCGTAGATTTCTCGTCAAAAAATC 780

Qy 781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCTGTTCTTCAAGAAATTTGATTTT 840
Db 781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCTGTTCTTCAAGAAATTTGATTTT 840

Qy 841 TTCAAAAAAATAGTTATTTTCTCTTATAAATAAGAAAACACTAGAAAATAGAGT 900
Db 841 TTCAAAAAAATAGTTATTTTCTCTTATAAATAAGAAAACACTAGAAAATAGAGT 900

Qy 901 TGCAGAGTAGCCCTAGAAATGTTTCCCAATAAATTAATCAATCACTGTATTAATTTTG 960
Db 901 TGCAGAGTAGCCCTAGAAATGTTTCCCAATAAATTAATCAATCACTGTATTAATTTTG 960

Qy 961 GCCAGCCCATTAATTAATTTAAACCGAACTGAAATGACGAAACCAAACTGACGTAT 1020
Db 961 GCCAGCCCATTAATTAATTTAAACCGAACTGAAATGACGAAACCAAACTGACGTAT 1020

```

```

Db 961 GCCAGCCCATTAATTAATTTAAACCGAACTGAAATGACGAAACCAAACTGAGCTAT 1020
Qy 1021 TTCTCTAGATTTAGTAAAAAGGAGAGAGAGAAAGAAATCTGTTTAAGTCAATGTCCC 1080
Db 1021 TTCTCTAGATTTAGTAAAAAGGAGAGAGAGAAAGAAATCTGTTTAAGTCAATGTCCC 1080

Qy 1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCTACGTCA 1140
Db 1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCTACGTCA 1140

Qy 1141 GGTTCGGAGCTCTCGTGTATCTCATGACATGACATACATGCTTTGCAACGGTTGTC 1200
Db 1141 GGTTCGGAGCTCTCGTGTATCTCATGACATGACATACATGCTTTGCAACGGTTGTC 1200

Qy 1201 TTGTTCCATGCTCCAAAGCCTTCCATTTCTGAAACCAAGAGATTAATCTCCAAACAT 1260
Db 1201 TTGTTCCATGCTCCAAAGCCTTCCATTTCTGAAACCAAGAGATTAATCTCCAAACAT 1260

Qy 1261 CCATCTTACTCATGCACTTCCATGCAACACGACATATGTTTCTGAAACCAATCCATT 1320
Db 1261 CCATCTTACTCATGCACTTCCATGCAACACGACATATGTTTCTGAAACCAATCCATT 1320

Qy 1321 AAAGATCACAACAGCTAGCGCTTCCCGCTAGCTTCCCTCTCTCTGCGATCTTTT 1380
Db 1321 AAAGATCACAACAGCTAGCGCTTCCCGCTAGCTTCCCTCTCTCTGCGATCTTTT 1380

Qy 1381 CGTCCACACGCAATG 1394
Db 1381 CGTCCACACGCAATG 1394

```

```

RESULT 2
AAH76332
ID AAH76332 standard; DNA; 1394 BP.
XX
XX AAH76332;
XX
XX 29-OCT-2001 (first entry)
XX
DE 2. mays Ms45 male tissue-preferred regulatory region encoding DNA.
XX
XX Ms45; male tissue; regulatory region; transcription; male fertility;
XX hybrid seed; ds.
XX
OS Zea mays.
XX
PN WO200160997-A2.
XX
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albersen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
XX WPI; 2001-514772/56.
XX
DR A male tissue-preferred regulatory region comprising nucleotide sequences
XX essential for initiating transcription of the Ms45 gene useful for
XX mediating fertility in a male plant.
XX
PS Claim 4; Page 46; 50pp; English.
XX
XX The invention provides a male tissue-preferred regulatory region (I)
XX comprising nucleotide sequences essential for initiating transcription of
XX the Ms45 gene. A method of mediating male fertility in a plant is
XX provided that involves introducing an expression vector comprising a
XX promoter operably linked to (I) into a plant where the exogenous gene
XX impacts male fertility of the plant and (II) controls expression of the
XX exogenous gene. A method of producing hybrid seeds is also provided. The

```

CC present sequence represents a nucleic acid sequence encoding an Ms45 male
 CC -cisbe preferred regulatory region from Z. mays
 XX
 SQ Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;
 Query Match 100.0%; Score 1394; DB 5; Length 1394;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCATGCTCTCTATGAAAAAGATGATACATGTGTCTATATCCGTTTTCTTAGGGTCC 60
 Db 1 CCATGCTCTCTATGAAAAAGATGATACATGTGTCTATATCCGTTTTCTTAGGGTCC 60

QY 61 CTCTCTCTGCTTATTAATGATGATCGGGGTTTACAAAACTTTCACGGGTGCTATGAT 120
 Db 61 CTCTCTCTGCTTATTAATGATGATCGGGGTTTACAAAACTTTCACGGGTGCTATGAT 120

QY 121 CTGCATGTTCCACTTCTCCACCTCGGGTTGACATTTCTTGATGTGGGTGCTCCAT 180
 Db 121 CTGCATGTTCCACTTCTCCACCTCGGGTTGACATTTCTTGATGTGGGTGCTCCAT 180

QY 181 CTGACCGAGGCCCATGACACCTTTGAGGACACCATCAAGGCGCTTTCGATGGCCCA 240
 Db 181 CTGACCGAGGCCCATGACACCTTTGAGGACACCATCAAGGCGCTTTCGATGGCCCA 240

QY 241 CGAGACGTATCGGGTGTGTGATCCAGGGGATATATGTCCCAATGCTGACCTTAT 300
 Db 241 CGAGACGTATCGGGTGTGTGATCCAGGGGATATATGTCCCAATGCTGACCTTAT 300

QY 301 TTATTAATCTTGTGATATTAATTTTGGAAAAATACAACTTATCTTTGTGTA 360
 Db 301 TTATTAATCTTGTGATATTAATTTTGGAAAAATACAACTTATCTTTGTGTA 360

QY 361 GGGCTCAGCATGATTTTTCGTTAGGGCCAGAAAAAGCAGAGCAGCATGCTTAGTG 420
 Db 361 GGGCTCAGCATGATTTTTCGTTAGGGCCAGAAAAAGCAGAGCAGCATGCTTAGTG 420

QY 421 TCCACTATTGGCATACCCAGAACAAAGATTAAAAAAATACCAAGTATCTATCTACT 480
 Db 421 TCCACTATTGGCATACCCAGAACAAAGATTAAAAAAATACCAAGTATCTATCTACT 480

QY 481 CGAAACGTATCAATGATGTTTAAAGAAACATCATTTAAACAGCATCTCTTAAAAA 540
 Db 481 CGAAACGTATCAATGATGTTTAAAGAAACATCATTTAAACAGCATCTCTTAAAAA 540

QY 541 CAAGCATATTTGCAAGAGACAATATGTTACAGTTTCAAAACATCTAAGACGACAAA 600
 Db 541 CAAGCATATTTGCAAGAGACAATATGTTACAGTTTCAAAACATCTAAGACGACAAA 600

QY 601 TTATATCGAAAGGTAAAGCTATGACGTTCAAGTTTCTTTTCATTTCTGTTATTTGTT 660
 Db 601 TTATATCGAAAGGTAAAGCTATGACGTTCAAGTTTCTTTTCATTTCTGTTATTTGTT 660

QY 661 ATGTGTTTTATATACATTTCTCTTCAATAGAGTGTCTTTTCCGATTTTAA 720
 Db 661 ATGTGTTTTATATACATTTCTCTTCAATAGAGTGTCTTTTCCGATTTTAA 720

QY 721 ATGACTATAAGCTATTTTATATTAAGAGCAGCATGCTAGATTTCTGTTCAAAAATC 780
 Db 721 ATGACTATAAGCTATTTTATATTAAGAGCAGCATGCTAGATTTCTGTTCAAAAATC 780

QY 781 TTTCTGATTTTTTAAAGCTAGTTGGCAACCTGTTTCTTCAAGAAATTTGATTTT 840
 Db 781 TTTCTGATTTTTTAAAGCTAGTTGGCAACCTGTTTCTTCAAGAAATTTGATTTT 840

QY 841 TTCAAAAAAATAGTTATTTCTCTTTAATAAATAGAAACATCTAGAAAAATGAGT 900
 Db 841 TTCAAAAAAATAGTTATTTCTCTTTAATAAATAGAAACATCTAGAAAAATGAGT 900

QY 901 TGCAGAGTAGCCCTAGAAATGTTTTTCCCAATAAATACATCACTGTGTAATTAATTTTG 960
 Db 901 TGCAGAGTAGCCCTAGAAATGTTTTTCCCAATAAATACATCACTGTGTAATTAATTTTG 960

QY 961 GCCAGCCCATTAATTTATTTAAACGAAACTGAAATCGAGCCGAAACCAATCTGAGCTAT 1020
 Db 961 GCCAGCCCATTAATTTATTTAAACGAAACTGAAATCGAGCCGAAACCAATCTGAGCTAT 1020

QY 1021 TTCTCTAGATTGTAATAAAGGAGAGAGAGAAAGAAATCAGTTTAAAGTCAATGTCCTCC 1080
 Db 1021 TTCTCTAGATTGTAATAAAGGAGAGAGAGAAAGAAATCAGTTTAAAGTCAATGTCCTCC 1080

QY 1081 TGAGATGTGCGGTTTGGCAAGATGACACCGCTAATCATAGGTCATAGGTCCTACGTCA 1140
 Db 1081 TGAGATGTGCGGTTTGGCAAGATGACACCGCTAATCATAGGTCATAGGTCCTACGTCA 1140

QY 1141 GGTTCGGAGCTCTGTGTGATCTCAATGGAATCTACATGCTTGTTCACCGTTCGTC 1200
 Db 1141 GGTTCGGAGCTCTGTGTGATCTCAATGGAATCTACATGCTTGTTCACCGTTCGTC 1200

QY 1201 TTGTTTCATGCTGTCGAAGCTTGCCTTAATTTGAAACAAAGAGATACCTTACCAAAAT 1260
 Db 1201 TTGTTTCATGCTGTCGAAGCTTGCCTTAATTTGAAACAAAGAGATACCTTACCAAAAT 1260

QY 1261 CCATCTTACTCATGCAACTTTCATGCAACACGCAATATGTTTCTGAAACCAATTCATT 1320
 Db 1261 CCATCTTACTCATGCAACTTTCATGCAACACGCAATATGTTTCTGAAACCAATTCATT 1320

QY 1321 AAAAGATCAAAACAGCTAGGTTCTCCGCTACTTCCCTCTCTCTGCGCATCTTTT 1380
 Db 1321 AAAAGATCAAAACAGCTAGGTTCTCCGCTACTTCCCTCTCTCTGCGCATCTTTT 1380

QY 1381 CGTCCACACCATG 1394
 Db 1381 CGTCCACACCATG 1394

RESULT 3
 AAX07409
 ID AAX07409 standard; DNA; 1394 BP.

AC AAX07409;
 AC 08-JUN-1999 (first entry)
 DT
 DT 08-JUN-1999 (first entry)
 DE Zea mays Ms45 male tissue-preferred regulatory region.
 XX
 XX Ms45; male; tissue-preferred; regulatory region; plant cells;
 KW plant tissue; differentiated; hybrid seed; fertility; ss.
 OS
 OS Zea mays.
 PN
 PN W09859061-A1.
 XX
 XX 30-DEC-1998.
 PD
 PD 19-JUN-1998; 98WO-US012895.
 PF
 PF 23-JUN-1997; 97US-00880499.
 PR
 PR (PION-) PIONEER HI-BRED INT INC.
 PA
 PA Albreten MC, Fox TW, Garnaat CW, Huffman GA, Kendall TL;
 PI
 PI WPI; 1999-105628/09.
 DR
 DR New nucleic acid encoding a Ms45 male tissue-preferred regulatory region.
 XX
 XX PT - useful in mediating plant fertility, especially hybrid seed production.
 PT
 PT
 XX
 XX
 PS Claim 3; Page 23-24; 39pp; English.

The sequence is that encoding an Ms45 male tissue-preferred regulatory region. It may be used in the construction of a vector for a method of producing exogenous genes in a male tissue-preferred manner, which is useful in restoring or conferring fertility, such as in hybrid seed production. In conferring fertility, a monocol/dicot plant is transformed with the exogenous nucleotide sequence (a male sterility gene, preferably

CC Ms45), which encodes a product selected from auxins, roib and dipheria
 CC toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
 CC and infertile plants
 XX

SQ Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;

Query Match 94.0%; Score 1311; DB 2; Length 1394;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 CCAATGCTCTATGAAAAAGATGATACATGCTATATCCGTTTCTTAGGCTCC 60
DB 1 CCAATGCTCTATGAAAAAGATGATACATGCTATATCCGTTTCTTAGGCTCC 60
QY 61 CTTCCTCTGCTTATTTCTGATCTGAATCGGGGTTACAAAAACTTCCACGGGTGATGAT 120
DB 61 CTTCCTCTGCTTATTTATGATCTGAATCGGGGTTACAAAAACTTCCACGGGTGATGAT 120
QY 121 CTCATGTTTCACCTTCCCACTCGGCTTGCACTTTCTTGATGTCGGTGGTCCAT 180
DB 121 CTCATGTTTCACCTTCCCACTCGGCTTGCACTTTCTTGATGTCGGTGGTCCAT 180
QY 181 CTGACCGAGGCCCATCAGACACCTTTGGGACACCCATCAAGGACCTTTGATGGCCA 240
DB 181 CTGACCGAGGCCCATCAGACACCTTTGGGACACCCATCAAGGACCTTTGATGGCCA 240
QY 241 CGAGAGCTATCGGGTGGTGAATCCAGGGGATATATGTCGCCCAATGCTACCTATA 300
DB 241 CGAGAGCTATCGGGTGGTGAATCCAGGGGATATATGTCGCCCAATGCTACCTATA 300
QY 301 TTATATCTCTTATGATATTTAATTTTGAATAAATAAATACTTATCTTTGTTGTA 360
DB 301 TTATATCTCTTATGATATTTAATTTTGAATAAATAAATACTTATCTTTGTTGTA 360
QY 361 GGGCTTCAGCATATGATTTTGGCTTGAAGGCCCAAGATCGAGAGACCAAGCTCTAGTG 420
DB 361 GGGCTTCAGCATATGATTTTGGCTTGAAGGCCCAAGATCGAGAGACCAAGCTCTAGTG 420
QY 421 TCCACTATTTGGCACTACCGAACAAGATTAAATAAATAAAGTAATCAATCACT 480
DB 421 TCCACTATTTGGCACTACCGAACAAGATTAAATAAATAAAGTAATCAATCACT 480
QY 481 CGAAGCTATCATGTAATGTTTAAAGAAATCATATTAATAACCAAGATCCTCTTAAAAA 540
DB 481 CGAAGCTATCATGTAATGTTTAAAGAAATCATATTAATAACCAAGATCCTCTTAAAAA 540
QY 541 CAAGCATATTTGCAAAAGACAAATTAATGTTTACAGTTTACAACTCTAAGACCGACAA 600
DB 541 CAAGCATATTTGCAAAAGACAAATTAATGTTTACAGTTTACAACTCTAAGACCGACAA 600
QY 601 TTATATCGAAGGTAAGCTATGAGCTCAGATTTTCTTTTCAATCTTGTTATTTTGT 660
DB 601 TTATATCGAAGGTAAGCTATGAGCTCAGATTTTCTTTTCAATCTTGTTATTTTGT 660
QY 661 ATTGTTTTTATATACATTTTCTCTCTTACATAGAGTATTTCTTCGATTTTATAA 720
DB 661 ATTGTTTTTATATACATTTTCTCTCTTACATAGAGTATTTCTTCGATTTTATAA 720
QY 721 ATGACTATAAAGTCATTTTATATTAAGACACGATCTCTTAATTTCTCTTAAAAATC 780
DB 721 ATGACTATAAAGTCATTTTATATTAAGACACGATCTCTTAATTTCTCTTAAAAATC 780
QY 781 TTCTGATTTTTTAAAGCTAGTTTGGCAACCTGTTTCTTCAAGAAATTTGATTTT 840
DB 781 TTCTGATTTTTTAAAGCTAGTTTGGCAACCTGTTTCTTCAAGAAATTTGATTTT 840
QY 841 TTCAAAAAAATAGTTTATTTCTCTTAAATAAGAAACACTTAGAAAAATAGAGT 900
DB 841 TTCAAAAAAATAGTTTATTTCTCTTAAATAAGAAACACTTAGAAAAATAGAGT 900
QY 901 TGGCAGCTAGCCCTAGAAATGTTTCCCAATTAATTCATCTGCTATTAATTTTGG 960
DB 901 TGGCAGCTAGCCCTAGAAATGTTTCCCAATTAATTCATCTGCTATTAATTTTGG 960

```

```

QY 961 GCCAGCCCCATTAATTTTAAACCGAAATCGAATCGAAGCAAAATCTAGACTAT 1020
DB 961 GCCAGCCCCATTAATTTTAAACCGAAATCGAATCGAAGCAAAATCTAGACTAT 1020
QY 1021 TTCTAGATTTAGTAAAAAGAGAGAGAGAGAAATAGTTTAAAGTATTTGCC 1080
DB 1021 TTCTAGATTTAGTAAAAAGAGAGAGAGAGAAATAGTTTAAAGTATTTGCC 1080
QY 1081 TGAGATGTCGGTTTGGCAAGATAGCCACCGTAATCATAGCTCATAGGCTTCACTCA 1140
DB 1081 TGAGATGTCGGTTTGGCAAGATAGCCACCGTAATCATAGCTCATAGGCTTCACTCA 1140
QY 1141 GGTGGGAGCTCTGTCATCTGATCATGATGGAATGCTACTCATGCTTCAACGTTCCGTC 1200
DB 1141 GGTGGGAGCTCTGTCATCTGATCATGATGGAATGCTACTCATGCTTCAACGTTCCGTC 1200
QY 1201 TTGTTTCATCTGTCGAAGCTTCTCTATTTGAAACGAAGAGATTAATCTCCCAACAT 1260
DB 1201 TTGTTTCATCTGTCGAAGCTTCTCTATTTGAAACGAAGAGATTAATCTCCCAACAT 1260
QY 1261 CCATCTTACTCATGCACTTCCATGCAACACGCAATATGTTTCTGTAAC 1311
DB 1261 CCATCTTACTCATGCACTTCCATGCAACACGCAATATGTTTCTGTAAC 1311

```

RESULT 4

AAH76333
 ID AAH76333 standard; DNA; 1394 BP.

AAH76333;

DT 29-OCT-2001 (first entry)

DE Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.

KM Ms45; male tissue; regulatory region; transcription; male fertility;

KM hybrid seed; ds.

OS Zea mays.

PN WO200160997-A2.

PD 23-AUG-2001.

PF 13-FEB-2001; 2001WO-US004527.

PR 15-FEB-2000; 2000US-00504487.

PA (PION-) PIONEER HI-BRED INT INC.

PI Albrechtsen MC, Fox TW, Garmaat CW, Huffman G, Kendall TL;

DR WPI; 2001-514772/56.

PT A male tissue-preferred regulatory region comprising nucleotide sequences

PT essential for initiating transcription of the Ms45 gene useful for

PS mediating fertility in a male plant.

PS Claim 4; Page 47; 50pp; English.

CC The invention provides a male tissue-preferred regulatory region (I)

CC comprising nucleotide sequences essential for initiating transcription of

CC the Ms45 gene. A method of mediating male fertility in a plant is

CC provided that involves introducing an expression vector comprising a

CC promoter operably linked to (I) into a plant where the exogenous gene

CC impacts male fertility of the plant and (II) controls expression of the

CC exogenous gene. A method of producing hybrid seeds is also provided. The

CC present sequence represents a nucleic acid sequence encoding an Ms45 male

CC tissue preferred regulatory region from Z. mays

XX Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;

Query Match	94.0%	Score 1311	DB 5	Length 1394	
Local Similarity	100.0%	Pred. No. 0			
Matches 1311	Conservative 0	Mismatches 0	Indels 0	Gaps 0	
Qy	1	CCATGCTCTCTATGAAAAAGATAGTACATGCTGTATATCCGTTTCTTAGGGGCC	60		
Db	1	CCATGCTCTCTATGAAAAAGATAGTACATGCTGTATATCCGTTTCTTAGGGGCC	60		
Qy	61	CTTCTTTCGCTTATTTACGTGACGTAAATCGGGGTTCACAAAAACTTCCACGGGTGCATGAT	120		
Db	61	CTTCTTTCGCTTATTTACGTGACGTAAATCGGGGTTCACAAAAACTTCCACGGGTGCATGAT	120		
Qy	121	CTCCATGTTCCACTTCTCCCACTTCGCGTTGCACATTTCTTGATGTCGGTGGTTCCCAT	180		
Db	121	CTCCATGTTCCACTTCTCCCACTTCGCGTTGCACATTTCTTGATGTCGGTGGTTCCCAT	180		
Qy	181	CTGACCGGAGGCCCATCAGACACTTTGGGGACACCCATCAAGGGGCTTTGGGATGGCCCA	240		
Db	181	CTGACCGGAGGCCCATCAGACACTTTGGGGACACCCATCAAGGGGCTTTGGGATGGCCCA	240		
Qy	241	CGAGACGTATCGGGTCGTGTGATCCAGGGATATATGTCCCCCAATCGTCACTATA	300		
Db	241	CGAGACGTATCGGGTCGTGTGATCCAGGGATATATGTCCCCCAATCGTCACTATA	300		
Qy	301	TTATTAATCTTTAGATATATTAATTAATTTTGGAAAAATTAACAAACTATATACCTTTGTGTA	360		
Db	301	TTATTAATCTTTAGATATATTAATTAATTTTGGAAAAATTAACAAACTATATACCTTTGTGTA	360		
Qy	361	GGGCTCAGCATAGATTTTTCGCTTAGGGCCCAAGAAATCGGAGACCAAGCTGTCTAGTG	420		
Db	361	GGGCTCAGCATAGATTTTTCGCTTAGGGCCCAAGAAATCGGAGACCAAGCTGTCTAGTG	420		
Qy	421	TCCACTATTTGGCACTACCCAGAACAAAGTTTAAAAAAATTAACAAAGTACATACTCACT	480		
Db	421	TCCACTATTTGGCACTACCCAGAACAAAGTTTAAAAAAATTAACAAAGTACATACTCACT	480		
Qy	481	CGAAAGCTATCATGTATGTTTAAAGAAACATCTATTAAAAACAAGATCCTTTAAAAA	540		
Db	481	CGAAAGCTATCATGTATGTTTAAAGAAACATCTATTAAAAACAAGATCCTTTAAAAA	540		
Qy	541	CAACCATATTTCGAAAGAGACAAATATAGTTACAGTTTACAAACATCTAAGAGGACAA	600		
Db	541	CAACCATATTTCGAAAGAGACAAATATAGTTACAGTTTACAAACATCTAAGAGGACAA	600		
Qy	601	TTATATCGAAAGGTAGCTATGACGTTTCTTTCATCTCTGTATTTTGT	660		
Db	601	TTATATCGAAAGGTAGCTATGACGTTTCTTTCATCTCTGTATTTTGT	660		
Qy	661	ATTGTTTTTATATACATTTTCTTCTTACATAGAGTATTTTCTCCGATTTTATATAA	720		
Db	661	ATTGTTTTTATATACATTTTCTTCTTACATAGAGTATTTTCTCCGATTTTATATAA	720		
Qy	721	ATGACTATAAAGTCATTTTATATPAGAGCAGCATGTCGTAAATTCCTGTTCAAAATC	780		
Db	721	ATGACTATAAAGTCATTTTATATPAGAGCAGCATGTCGTAAATTCCTGTTCAAAATC	780		
Qy	781	TTTCTGATTTTTTTPAAGAGTATGTTTGGCAACCCGTTCCTTCAAGAAATTTTGATTT	840		
Db	781	TTTCTGATTTTTTTPAAGAGTATGTTTGGCAACCCGTTCCTTCAAGAAATTTTGATTT	840		
Qy	841	TTCAAAAAAATTAAGTTTATTTCTCTTATPAAAAATAGAAAAACATTAGAAAAATAGAGT	900		
Db	841	TTCAAAAAAATTAAGTTTATTTCTCTTATPAAAAATAGAAAAACATTAGAAAAATAGAGT	900		
Qy	901	TGCCAGACTAGCCCTAGAAATGTTTCCCAATPAAATTACATCACTGTGTATATATTGG	960		
Db	901	TGCCAGACTAGCCCTAGAAATGTTTCCCAATPAAATTACATCACTGTGTATATATTGG	960		
Qy	961	GCAGAGCCCATPAAATTAATTTAAACCGAAATCGAAGGAAACCAAAATCTAGAGTAT	1020		
Db	961	GCAGAGCCCATPAAATTAATTTAAACCGAAATCGAAGGAAACCAAAATCTAGAGTAT	1020		
Qy	1021	TTCTCTAGATTAGTAAAAAGGAGAGAGAGAAATATGATCATGTGTCC	1080		

ID	Sequence	Length
Db	1021 TTCTTCAATTAAAGTAAAAAGGAGAGAGAAAGAAATCAAGTTTAAATCAATTCCTCC	108
Qy	1081 TGAGATGAGCGGTTGGCAAGATAGCCAGCATATCA TAGCTCA TAGGTCCTAACGTC	114
Db	1081 TGAGATGAGCGGTTGGCAAGATAGCCAGCATATCA TAGCTCA TAGGTCCTAACGTC	114
Qy	1141 GGTTCGAGCGCTCTGTCATCTTCACATGGCATACTAAGCTTGTTCACACGTTGCTC	120
Db	1141 GGTTCGAGCGCTCTGTCATCTTCACATGGCATACTAAGCTTGTTCACACGTTGCTC	120
Qy	1201 TTGTTCCATCGTCCAAAGCCTTGCTATTTGAAACCAAGAGATACCTAATCCCAACAT	126
Db	1201 TTGTTCCATCGTCCAAAGCCTTGCTATTTGAAACCAAGAGATACCTAATCCCAACAT	126
Qy	1261 CCATCTTACCTCATGCACTTCATCAGCAACAGCAACGCAATATGTTCTCGAAC	131
Db	1261 CCATCTTACCTCATGCACTTCATCAGCAACAGCAACGCAATATGTTCTCGAAC	131
RESULT 5		
AAH76340	AAH76340 standard; DNA; 255 BP.	
XX	AAH76340;	
XX	29-OCT-2001 (first entry)	
DE	Z. mays Ms45 promoter fragment.	
XX	Ms45; male tissue; regulatory region; transcription; male fertility;	
KM	hybrid seed; promoter; ds.	
XX	Zea mays.	
OS	WO200160997-A2.	
XX	23-AUG-2001.	
XX	13-FEB-2001; 2001WO-US004527.	
PF	15-FEB-2000; 2000US-00504487.	
XX	(PION-) PIONEER HI-BRED INT INC.	
PA	Albertsen MC, Fox TW, Garnaat CW, Hufman G, Kendall TL;	
XX	WPI; 2001-514772/56.	
DR		
XX		
PT	A male tissue-preferred regulatory region comprising nucleotide sequences	
PT	essential for initiating transcription of the Ms45 gene useful for	
PT	mediating fertility in a male plant.	
XX		
PS	Example 5; Fig 8; 50pp; English.	
XX		
CC	The invention provides a male tissue-preferred regulatory region (I)	
CC	comprising nucleotide sequences essential for initiating transcription of	
CC	the Ms45 gene. A method of mediating male fertility in a plant is	
CC	provided that involves introducing an expression vector comprising a	
CC	promoter operably linked to (I) into a plant where the exogenous gene	
CC	impacts male fertility of the plant and (I) controls expression of the	
CC	exogenous gene. A method of producing hybrid seeds is also provided. The	
CC	present sequence represents a Z. mays Ms45 promoter fragment	
XX		
SO	Sequence 255 BP; 59 A; 86 C; 39 G; 71 T; 0 U; 0 Other;	
Query Match 11.3%; Score 157; DB 5; Length 255;		
Best Local Similarity 100.0%; Pred. No. 9.9e-65;		
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy	1155 CGGTGATCTCAACATGCGATACATACATGCTTGTTCACACCGTTCGCTTGTTCATCGTCC	1214
Db	15 CGGTGATCTCAACATGCGATACATACATGCTTGTTCACACCGTTCGCTTGTTCATCGTCC	74

```
QY 1215 AAGCCTTGCTATTCTGAACCAAGAGATACCTACTCCGAACATCCATTACTGAC 1274
DB 75 AAGCCTTGCTATTCTGAACCAAGAGATACCTACTCCGAACATCCATTACTGAC 134
QY 1275 CAATCTTCATGCAACACGACATATGTTCTCTGAAC 1311
DB 135 CAATCTTCATGCAACACGACATATGTTCTCTGAAC 171

RESULT 6
AAH76334
ID AAH76334 standard; DNA; 158 BP.
XX
XX AAH76334;
AC
XX
XX 29-OCT-2001 (first entry) .
DT
XX
XX 2. may5 Ms45 male tissue-preferred regulatory region fragment.
DE
XX
XX Ms45; male tissue; regulatory region; transcription; male fertility;
KM hybrid seed; ds.
XX
XX Zea mays.
OS
XX WO200160997-A2.
PN
XX 23-AUG-2001.
PD
XX 13-FEB-2001; 2001WO-US004527.
PF
XX 15-FEB-2000; 2000US-00504487.
PR
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX WPI; 2001-514772/56.
DR
XX
XX A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the Ms45 gene useful for
PT mediating fertility in a male plant.
XX
XX Claim 5; Page 47; 50pp; English.
XX
XX The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the Ms45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a DNA fragment -38 to -195 bases upstream of
CC the TATA box of a Z. may5 Ms45 male-tissue preferred regulatory region
CC nucleotide sequence
XX
XX
SQ Sequence 158 BP; 41 A; 50 C; 21 G; 46 T; 0 U; 0 Other;

Query Match 8.0%; Score 111; DB 5; Length 158;
Best Local Similarity 100.0%; Pred. No. 1.3e-42;
Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1201 TTGTTCCATGATCAAGCCTTGCTATTCTGAACCAAGAGATACCTACTCCGAACAT 1260
DB 48 TTGTTCCATGATCAAGCCTTGCTATTCTGAACCAAGAGATACCTACTCCGAACAT 107
QY 1261 CCATCTTACTCATGCAACTTCATGCAACACGACATATGTTCTCTGAAC 1311
DB 108 CCATCTTACTCATGCAACTTCATGCAACACGACATATGTTCTCTGAAC 158

RESULT 7
AAH76336
```

```
ID AAH76336 standard; DNA; 50 BP.
XX
XX AAH76336;
AC
XX
XX 29-OCT-2001 (first entry)
DT
XX
XX 2. may5 Ms45 male tissue-preferred regulatory region fragment.
DE
XX
XX Ms45; male tissue; regulatory region; transcription; male fertility;
KM hybrid seed; ds.
XX
XX Zea mays.
OS
XX WO200160997-A2.
PN
XX 23-AUG-2001.
PD
XX 13-FEB-2001; 2001WO-US004527.
PF
XX 15-FEB-2000; 2000US-00504487.
PR
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX WPI; 2001-514772/56.
DR
XX
XX A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the Ms45 gene useful for
PT mediating fertility in a male plant.
XX
XX Claim 5; Page 47; 50pp; English.
XX
XX The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the Ms45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a DNA fragment -72 to -111 bases upstream of
CC the TATA box of a Z. may5 Ms45 male-tissue preferred regulatory region
CC nucleotide sequence
XX
XX
SQ Sequence 50 BP; 18 A; 17 C; 5 G; 10 T; 0 U; 0 Other;

Query Match 3.6%; Score 50; DB 5; Length 50;
Best Local Similarity 100.0%; Pred. No. 2.6e-13;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1229 CTGAACCAAGAGATACCTACTCCGAACATCCATCTTACTCATGCAAC 1278
DB 1 CTGAACCAAGAGATACCTACTCCGAACATCCATCTTACTCATGCAAC 50

RESULT 8
AAH76337
ID AAH76337 standard; DNA; 40 BP.
XX
XX AAH76337;
AC
XX
XX 29-OCT-2001 (first entry)
DT
XX
XX 2. may5 Ms45 male tissue-preferred regulatory region fragment.
DE
XX
XX Ms45; male tissue; regulatory region; transcription; male fertility;
KM hybrid seed; ds.
XX
XX Zea mays.
OS
XX WO200160997-A2.
PN
XX 23-AUG-2001.
PD
```

XX 13-FEB-2001; 2001WO-US004527.
PF 15-FEB-2000; 2000US-00504487.
PR (PION-) PIONEER HI-BRED INT INC.
XX
XX PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
DR WPI; 2001-514772/56.
XX
XX A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the MS45 gene useful for
PT mediating fertility in a male plant.
XX
XX Claim 14; Page 32; 50pp; English.
XX
XX The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the MS45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a DNA fragment upstream of the TATA box of a
CC Z. mays Ms45 male-tissue preferred regulatory region nucleotide sequence
XX
XX Sequence 40 BP; 14 A; 14 C; 3 G; 9 T; 0 U; 0 Other;
SQ
XX
XX Query Match 2.9%; Score 40; DB 5; Length 40;
Best Local Similarity 100.0%; Pred. No. 1.6e-08;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1239 AGGATACCTACTCCCAACATCATCTTACTCATGCAC 1278
DB 1 AGGATACCTACTCCCAACATCATCTTACTCATGCAC 40
XX
XX RESULT 9
AAH76335
ID AAH76335 standard; DNA; 30 BP.
XX
XX AAH76335;
AC
XX
XX 29-OCT-2001 (first entry)
DT
XX
XX Z. mays Ms45 male tissue-preferred regulatory region fragment.
DE
XX
XX Ms45; male tissue; regulatory region; transcription; male fertility;
KM hybrid seed; db.
XX
XX Zea mays.
OS
XX
XX WO200160997-A2.
PN
XX
XX 23-AUG-2001.
PD
XX
XX 13-FEB-2001; 2001WO-US004527.
PF
XX
XX 15-FEB-2000; 2000US-00504487.
PR
XX
XX (PION-) PIONEER HI-BRED INT INC.
PA
XX
XX PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
PI
XX
XX WPI; 2001-514772/56.
DR
XX
XX A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the MS45 gene useful for
PT mediating fertility in a male plant.
XX
XX Claim 5; Page 47; 50pp; English.
XX

CC The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the MS45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a DNA fragment -152 to -181 bases upstream of
CC the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
XX
XX nucleotide sequence
SQ
XX
XX Sequence 30 BP; 4 A; 9 C; 5 G; 12 T; 0 U; 0 Other;
SQ
XX
XX Query Match 2.2%; Score 30; DB 5; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1179 CATGCTGTTCACCGTGTCTGTTC 1208
DB 1 CATGCTGTTCACCGTGTCTGTTC 30
XX
XX RESULT 10
ABL40437/C
ID ABL40437 standard; DNA; 849 BP.
XX
XX ABL40437;
AC
XX
XX 10-JUN-2002 (first entry)
DT
XX
XX Sunflower chitinase promoter sequence.
DE
XX
XX Sunflower; chitinase; lipid transfer protein; LTP; transcription;
KM sclerotinia; genetic engineering; disease resistance; insecticide;
KM antifungal; viricide; nematocide; antimicrobial; antibacterial;
KM gene therapy; gene; promoter; ds.
XX
XX Helianthus annuus.
OS
XX
XX Key Location/Qualifiers
FH
XX
XX CAAAT_signal 723..726
FT
XX
XX TATA_signal 807..811
FT
XX
XX /*tag= b
XX
XX WO200214502-A2.
PN
XX
XX 21-FEB-2002.
PD
XX
XX 08-AUG-2001; 2001WO-US041629.
PF
XX
XX 11-AUG-2000; 2000US-0224603P.
PR
XX
XX 07-AUG-2001; 2001US-00923844.
PR
XX
XX (PION-) PIONEER HI-BRED INT INC.
PA
XX
XX PI Bao Z, Lu G;
PI
XX
XX WPI; 2002-269194/31.
DR
XX
XX New genes and promoters from Helianthus annuus, for genetically
PT manipulating plants to enhance disease resistance or resistance to
PT pathogen, and for producing proteins for controlling plant diseases
PT caused by these pathogens.
PT
XX
XX Claim 1; Fig 4; 89pp; English.
PS
XX
XX The invention relates to sclerotinia-inducible genes and promoters capable
CC of initiating transcription from sunflower. The polynucleotides are
CC useful for genetically manipulating plants to enhance disease resistance,
CC or resistance to insect, fungal pathogens, viruses or nematodes. The
CC nucleotide sequences are also useful as genetic markers in disease-
CC resistance breeding programs. The polynucleotides are also useful for

CC producing proteins useful in compositions. The compositions are useful
CC reducing or eliminating damage to plants caused by plant pathogens. The
CC compositions find use in agricultural and pharmaceutical compositions as
CC antifungal and antimicrobial agents. For agricultural purposes, the
CC compositions may be used in sprays for control of plant disease. As
CC pharmaceutical compositions, the agents are useful as antibacterial and
CC antimicrobial treatments. The present sequence represents the promoter
CC fragment of the sunflower chitinase gene
XX
SQ Sequence 849 BP; 295 A; 139 C; 135 G; 280 T; 0 U; 0 Other;
Query Match 1.6%; Score 23; DB 6; Length 849;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 854 AGTTATTTCTCTTATAAAT 876
DB 556 AGTTATTTCTCTTATAAAT 534
RESULT 11
ABV56505
ID ABV56505 standard; cDNA, 333 BP.
AC ABV56505;
XX
XX
DT 17-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 56496.
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KM pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-018319P.
XX
PR 16-MAR-2000; 2000US-0189662P.
XX
PR 25-MAY-2000; 2000US-0207454P.
XX
PR 09-JUN-2000; 2000US-0211314P.
XX
PR 18-JUL-2000; 2000US-0219007P.
XX
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX
PI Schlegel R, Endege WO, Monahan JR;
XX
XX WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
XX
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 10900; 11750pp; English.
XX
XX
XX The invention relates to an isolated nucleic acid molecule (1) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (1) is useful for: (a) assessing whether
XX a patient is afflicted with prostate cancer; (b) monitoring the efficacy
XX progression of prostate cancer in a patient; (c) assessing the efficacy
XX of a test compound to inhibit prostate cancer in a patient; (d) assessing
XX the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX (e) selecting a composition for inhibiting prostate cancer in a patient;
XX (f) assessing the prostate cell carcinogenic potential of a compound; (g)
XX determining whether prostate cancer has metastasized in a patient; (h)
XX assessing the aggressiveness or indolence of prostate cancer in a patient
XX ; (1) is also useful as a pharmacodynamic or pharmacogenomic marker
XX

SQ Sequence 333 BP; 75 A; 69 C; 46 G; 143 T; 0 U; 0 Other;
Query Match 1.5%; Score 21; DB 5; Length 333;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 93 TTACAAAAAATTCACGGGT 113
DB 89 TTACAAAAAATTCACGGGT 109
RESULT 12
ABA58867/c
ID ABA58867 standard; DNA; 459 BP.
XX
AC ABA58867;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human foetal liver single exon nucleic acid probe #7172.
XX
XX Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000669.
XX
PR 04-FEB-2000; 2000US-0180312P.
XX
PR 26-MAY-2000; 2000US-0207456P.
XX
PR 30-JUN-2000; 2000US-0060840P.
XX
PR 03-AUG-2000; 2000US-0063236P.
XX
PR 21-SEP-2000; 2000US-0234687P.
XX
PR 27-SEP-2000; 2000US-0236359P.
XX
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483447/52.
XX
DR
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX
XX gene expression in human foetal liver.
XX
PS Claim 1; SEQ ID NO 7172; 639pp + Sequence Listing; English.
XX
XX
XX The invention relates to a single exon nucleic acid probe for measuring
XX human gene expression in a sample derived from human foetal liver. The
XX single exon nucleic acid probes may be used for predicting, measuring and
XX displaying gene expression in samples derived from human foetal liver. The
XX present sequence is a single exon nucleic acid probe of the invention.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 459 BP; 86 A; 149 C; 70 G; 154 T; 0 U; 0 Other;
Query Match 1.5%; Score 21; DB 4; Length 459;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1037 AAAGGAGAGAGAGAGAGA 1057
DB 118 AAAGGAGAGAGAGAGAGA 98
RESULT 13
AA138581/c
ID AA138581 standard; DNA; 459 BP.


```

XX AC AAI38581;
XX
XX DT 17-OCT-2001 (first entry)
XX
XX DE Probe #7267 used to measure gene expression in human placenta sample.
XX
XX KW Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder; ss.
XX
XX OS Homo sapiens.
XX
XX PN WO200157272-A2.
XX
XX PD 09-AUG-2001.
XX
XX PF 30-JAN-2001; 2001WO-US000663.
XX
XX PR 04-FEB-2000; 2000US-0180312P.
XX
XX PR 26-MAY-2000; 2000US-0207456P.
XX
XX PR 30-JUN-2000; 2000US-0207456P.
XX
XX PR 03-AUG-2000; 2000US-00608408.
XX
XX PR 21-SEP-2000; 2000US-0234687P.
XX
XX PR 27-SEP-2000; 2000US-0236359P.
XX
XX PR 04-OCT-2000; 2000GB-00024263.
XX
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX DR WPI; 2001-488907/53.
XX
XX PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human placenta.
XX
XX PS Claim 25; SEQ ID NO 7267; 654bp; English.
XX
XX CC The present invention relates to single exon nucleic acid probes (SENP).
XX CC The present sequence is one such probe. The probes are useful for
XX CC producing a microarray for predicting, measuring and displaying gene
XX CC expression in samples derived from human placenta. The probes are useful
XX CC for antenatal diagnosis of human genetic disorders
XX
XX SO Sequence 459 BP; 86 A; 149 C; 70 G; 154 T; 0 U; 0 Other;

Query Match      1.5%; Score 21; DB 4; Length 459;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1037 AAAGGAGAGAGAGAGAGA 1057
DB      118 AAAGGAGAGAGAGAGAGA 98

RESULT 14
AAK32768/c
ID AAK32768 standard; DNA; 459 BP.
AC
XX AAK32768;
XX
XX DT 06-NOV-2001 (first entry)
XX
XX DE Human bone marrow expressed single exon probe SEQ ID NO: 7325.
XX
XX KW Human; bone marrow expressed exon; gene expression analysis; probe;
XX microarray; cancer; leukemia; lymphoma; myeloma; ss.
XX
XX OS Homo sapiens.
XX
XX PN WO200157276-A2.
XX
XX PD 09-AUG-2001.
XX

```

```

PF 30-JAN-2001; 2001WO-US000668.
XX
XX PR 04-FEB-2000; 2000US-0180312P.
XX
XX PR 26-MAY-2000; 2000US-0207456P.
XX
XX PR 30-JUN-2000; 2000US-00608408.
XX
XX PR 03-AUG-2000; 2000US-0234687P.
XX
XX PR 21-SEP-2000; 2000US-0236359P.
XX
XX PR 27-SEP-2000; 2000US-0236359P.
XX
XX PR 04-OCT-2000; 2000GB-00024263.
XX
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX DR WPI; 2001-488900/53.
XX
XX PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human bone marrow.
XX
XX PS Example 4; SEQ ID NO 7325; 658bp + Sequence Listing; English.
XX
XX CC The present invention provides a number of single exon nucleic acid
XX CC probes which are derived from genomic sequences expressed in the human
XX CC bone marrow. They can be used to measure gene expression in bone marrow
XX CC samples, which may enable the improved diagnosis and treatment of cancers
XX CC such as lymphoma, leukemia and myeloma. The present sequence is one of
XX CC the probes of the invention
XX
XX SO Sequence 459 BP; 86 A; 149 C; 70 G; 154 T; 0 U; 0 Other;

Query Match      1.5%; Score 21; DB 4; Length 459;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1037 AAAGGAGAGAGAGAGAGA 1057
DB      118 AAAGGAGAGAGAGAGAGA 98

RESULT 15
AAK07027/c
ID AAK07027 standard; DNA; 459 BP.
AC AAK07027;
XX
XX DT 05-NOV-2001 (first entry)
XX
XX DE Human brain expressed single exon probe SEQ ID NO: 7018.
XX
XX KW Human; brain expressed exon; gene expression analysis; probe; microarray;
XX Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
XX ss.
XX
XX OS Homo sapiens.
XX
XX PN WO200157275-A2.
XX
XX PD 09-AUG-2001.
XX
XX PF 30-JAN-2001; 2001WO-US000667.
XX
XX PR 04-FEB-2000; 2000US-0180312P.
XX
XX PR 26-MAY-2000; 2000US-0207456P.
XX
XX PR 30-JUN-2000; 2000US-00608408.
XX
XX PR 03-AUG-2000; 2000US-0234687P.
XX
XX PR 21-SEP-2000; 2000US-0236359P.
XX
XX PR 27-SEP-2000; 2000US-0236359P.
XX
XX PR 04-OCT-2000; 2000GB-00024263.
XX
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX

```

DR WPI; 2001-483446/52.

XX Single exon nucleic acid probes for analyzing gene expression in human
PT brains.

XX PS Example 4; SEQ ID NO 7018; 650pp + Sequence listing; English.

XX CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention

XX SQ Sequence 459 BP; 86 A; 149 C; 70 G; 154 T; 0 U; 0 Other;

Query Match 1.5%; Score 21; DB 4; Length 459;

Best Local Similarity 100.0%; Pred. No. 21;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1037 AAAGGAGAGAGAGAGAGA 1057

Db 118 AAAGGAGAGAGAGAGAGA 98

Search completed: September 16, 2005, 13:08:39
Job time : 584.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using bw model

Run on: September 15, 2005, 19:42:02 ; Search time 176.5 Seconds
(without alignments)
12923.341 Million cell updates/sec

Title: US-10-713-381-1

Perfect score: 1394
Sequence: 1 cccatgctgcctcctcgaagaaa.....cttttcgccaccaccaccatg 1394

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1202784 seqs, 818138359 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B COMB.seq: *
3: /cgn2_6/prodata/1/ina/6A COMB.seq: *
4: /cgn2_6/prodata/1/ina/6B COMB.seq: *
5: /cgn2_6/prodata/1/ina/6C COMB.seq: *
6: /cgn2_6/prodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1394	100.0	1394	3	US-08-880-499-1
2	1311	94.0	1394	4	US-08-880-499-2
3	21	1.5	50263	4	US-09-949-016-13563
4	20	1.4	601	4	US-09-949-016-124481
5	20	1.4	601	4	US-09-949-016-147075
6	20	1.4	2523	2	US-08-410-784A-3
7	20	1.4	21679	4	US-09-949-016-15250
8	20	1.4	65424	4	US-09-949-016-12426
9	20	1.4	374159	4	US-09-949-016-15868
10	19	1.4	601	4	US-09-949-016-17933
11	19	1.4	601	4	US-09-949-016-17934
12	19	1.4	601	4	US-09-949-016-41823
13	19	1.4	601	4	US-09-949-016-41824
14	19	1.4	601	4	US-09-949-016-145344
15	19	1.4	601	4	US-09-949-016-145345
16	19	1.4	766	4	US-09-328-352-2409
17	19	1.4	1584	1	US-08-247-808A-10
18	19	1.4	1584	1	US-08-453-942-10
19	19	1.4	1584	1	US-08-926-885A-10
20	19	1.4	1584	5	PCT-US94-05290-10
21	19	1.4	1833	1	US-08-365-470-2
22	19	1.4	3834	3	US-09-209-668-18
23	19	1.4	3834	4	US-09-949-016-71
24	19	1.4	3834	1	US-08-365-470-1
25	19	1.4	3858	2	US-08-344-155C-98
26	19	1.4	3858	3	US-09-009-490A-88
27	19	1.4	3862	4	US-09-949-016-4401

c 28	19	1.4	3863	3	US-08-482-073-1	Sequence 1, Appl1
c 29	19	1.4	3863	6	5217870-1	Patent No. 5217870
c 30	19	1.4	3863	6	5217870-1	Patent No. 5217870
c 31	19	1.4	9370	1	US-08-320-559-27	Sequence 27, Appl1
c 32	19	1.4	9370	3	US-08-545-860D-27	Sequence 27, Appl1
c 33	19	1.4	9370	5	PCT-US94-04496-27	Sequence 27, Appl1
c 34	19	1.4	9391	1	US-08-320-559-25	Sequence 25, Appl1
c 35	19	1.4	9391	3	US-08-545-860D-25	Sequence 25, Appl1
c 36	19	1.4	9391	5	PCT-US94-04496-25	Sequence 25, Appl1
c 37	19	1.4	15384	4	US-09-949-016-11813	Sequence 11813, A
c 38	19	1.4	15413	4	US-09-949-016-16143	Sequence 16143, A
c 39	19	1.4	71645	4	US-09-949-016-12126	Sequence 12126, A
c 40	19	1.4	71651	4	US-09-949-016-17258	Sequence 17258, A
c 41	19	1.4	105189	4	US-09-949-016-13029	Sequence 13029, A
c 42	19	1.4	133559	4	US-09-949-016-15845	Sequence 15845, A
c 43	19	1.4	144158	4	US-09-949-016-11755	Sequence 11755, A
c 44	19	1.4	144158	4	US-09-949-016-12936	Sequence 12936, A
c 45	19	1.4	228896	4	US-09-949-016-17127	Sequence 17127, A

ALIGNMENTS

RESULT 1
US-08-880-499-1

; Sequence 1, Application US/08880499
; Patent No. 6037523

; GENERAL INFORMATION:

; APPLICANT: Alberson, Marc C.

; APPLICANT: Fox, Tim W.

; APPLICANT: Carl, Gary A.

; APPLICANT: Kendall, Timmy L.

; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESSES:

; ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.

; STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.

; CITY: Johnston

; STATE: Iowa

; COUNTRY: USA

; ZIP: 50131

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/880,499

; FILING DATE: CONCURRENTLY HERewith

; CLASSIFICATION: 800

; ATTORNEY/AGENT INFORMATION:

; NAME: Sweeney, Patricia A.

; REGISTRATION NUMBER: 32,733

; REFERENCE/DOCKET NUMBER: 0578

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (515) 248-4800

; TELEFAX: (515) 248-4844

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1394 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; US-08-880-499-1

Query Match 100.0%; Score 1394; DB 3; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCATGCTGTCTATGAAAAAGATGATACATGTCATATCCGTTTCTTAGGGTCC 60
Db 1 CCATGCTGTCTATGAAAAAGATGATACATGTCATATCCGTTTCTTAGGGTCC 60
QY 61 CTTCTTCTGCTTATTACTGACTGAAATCGGGGTTTCAAAAACTTCCACGGGTGCATGAT 120
Db 61 CTTCTTCTGCTTATTACTGACTGAAATCGGGGTTTCAAAAACTTCCACGGGTGCATGAT 120
QY 121 CTCATGTTCCATCTTCCACCTCGGCTGCACTTCTTGGATGCGGTGCTCCCAT 180
Db 121 CTCATGTTCCATCTTCCACCTCGGCTGCACTTCTTGGATGCGGTGCTCCCAT 180
QY 181 CTGACCGAGGCCCATGACACCTTTCGGGACACCCATCAAGGSCCTTTCGATGAGCCCA 240
Db 181 CTGACCGAGGCCCATGACACCTTTCGGGACACCCATCAAGGSCCTTTCGATGAGCCCA 240
QY 241 CGAGAGCTATCGGGTGTGTGATCCAGGGGATATATGTCCTCCCAACATGTCACCTATA 300
Db 241 CGAGAGCTATCGGGTGTGTGATCCAGGGGATATATGTCCTCCCAACATGTCACCTATA 300
QY 301 TTATTTATCTTTAGATTATTTAATTTTGGAAAAATACAAACTTATCTTTTGCTA 360
Db 301 TTATTTATCTTTAGATTATTTAATTTTGGAAAAATACAAACTTATCTTTTGCTA 360
QY 361 GGGCTTCAGCATGATTTTCGCTTAGGGCCAGAAAATGCGAGGACGAGCCATGCTAGTG 420
Db 361 GGGCTTCAGCATGATTTTCGCTTAGGGCCAGAAAATGCGAGGACGAGCCATGCTAGTG 420
QY 421 TCCACTTATGGCACTACCCAGAACAGATTTAAAAAATACCAAGTAACTATATCCACT 480
Db 421 TCCACTTATGGCACTACCCAGAACAGATTTAAAAAATACCAAGTAACTATATCCACT 480
QY 481 CGAAGCTATCATGTATGTTTAAAGAAACATCTATTTAAACACAGATCCTCTTAAAAA 540
Db 481 CGAAGCTATCATGTATGTTTAAAGAAACATCTATTTAAACACAGATCCTCTTAAAAA 540
QY 541 CAAGCATATTTGCAAAAGACAAATATGTATACAGTTTACAAACATCTAAGACGACAA 600
Db 541 CAAGCATATTTGCAAAAGACAAATATGTATACAGTTTACAAACATCTAAGACGACAA 600
QY 601 TTATATCGAAAGGTAAAGCTATGACGTTGAGATTTTCTTTTCAATCTTGTTATTTGTT 660
Db 601 TTATATCGAAAGGTAAAGCTATGACGTTGAGATTTTCTTTTCAATCTTGTTATTTGTT 660
QY 661 ATGTTTTTATATACATTTCTTCTCTTATACATAGATGATTTTCTTCCGATTTTAAAA 720
Db 661 ATGTTTTTATATACATTTCTTCTCTTATACATAGATGATTTTCTTCCGATTTTAAAA 720
QY 721 ATGACTATAAAGTCAATTTTATATAAGAGACGCAATGCTAGATTTCTGTTCAAAAAATC 780
Db 721 ATGACTATAAAGTCAATTTTATATAAGAGACGCAATGCTAGATTTCTGTTCAAAAAATC 780
QY 781 TTTTCGATTTTAAAGCTAGTTTGGCAACCTGTTTCTTCAAGAAATTTTGATTTT 840
Db 781 TTTTCGATTTTAAAGCTAGTTTGGCAACCTGTTTCTTCAAGAAATTTTGATTTT 840
QY 841 TTCAAAAAAATAGTTTATTTCTCTTATATAAATAGAAAAACATAGAAAAATAGAGT 900
Db 841 TTCAAAAAAATAGTTTATTTCTCTTATATAAATAGAAAAACATTAAGAAAAATAGAGT 900
QY 901 TGGCAGACTAGCCCTAGAAATGTTTCCCAATAAATTAACATCACTGTATATAATTTTGG 960
Db 901 TGGCAGACTAGCCCTAGAAATGTTTCCCAATAAATTAACATCACTGTATATAATTTTGG 960
QY 961 GCCAGCCCCATTAATTTTAAACCGAAATCGAAATCGAGCGAAACCAATCTGACCTAT 1020
Db 961 GCCAGCCCCATTAATTTTAAACCGAAATCGAAATCGAGCGAAACCAATCTGACCTAT 1020
QY 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAATCACTTTTAAAGCATTTGCC 1080
Db 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAATCACTTTTAAAGCATTTGCC 1080
QY 1081 TGAGATGTGCGGTTTGGCAACGATAGCCAGTAAATCATAGCTATAGGTGCTTACGTCA 1140

Db 1081 TGAGATGTGCGGTTTGGCAACGATAGCCAGCTAATACATAGCTATAGGTGCTTACGTCA 1140
QY 1141 GGTTCGGACGCTCTGTGTATCTCATGACATGCAATACATGTTGTTCAACCGTTGCTC 1200
Db 1141 GGTTCGGACGCTCTGTGTATCTCATGACATGCAATACATGTTGTTCAACCGTTGCTC 1200
QY 1201 TTGTTTCATGCTGCAAGCCTTGCTATTGTAACCAAGAGGATPACTACTCCCAACAT 1260
Db 1201 TTGTTTCATGCTGCAAGCCTTGCTATTGTAACCAAGAGATPACTACTCCCAACAT 1260
QY 1261 CCATCTTACTCATGCACTTTCATGCAACAGCAATATGTTTCTGAAACCAATCCATT 1320
Db 1261 CCATCTTACTCATGCACTTTCATGCAACAGCAATATGTTTCTGAAACCAATCCATT 1320
QY 1321 AAAGATCACAAAGCTAGGCTTCTCCGCTAGCTTCCCTCTCTCTGCGCATCTTTT 1380
Db 1321 AAAGATCACAAAGCTAGGCTTCTCCGCTAGCTTCCCTCTCTCTGCGCATCTTTT 1380
QY 1381 CGTCCACGACGATG 1394
Db 1381 CGTCCACGACGATG 1394

RESULT 2
US-08-880-499-2
Sequence 2, Application US/08880499
Patent No. 6037523
GENERAL INFORMATION:
APPLICANT: Albertson, Marc C.
APPLICANT: Fox, Tim W.
APPLICANT: Carl, Garnat W.
APPLICANT: Huffman, Gary A.
APPLICANT: Kendall, Timmy L.
TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION
NUMBER OF INVENTIONS: 2
NUMBER OF SEQUENCES: 2
TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION
CORRESPONDENCE ADDRESS:
ADDRESSER: PIONEER HI-BRED INTERNATIONAL, INC.
STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.
CITY: Johnston
STATE: Iowa
COUNTRY: USA
ZIP: 50131
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/880 499
FILING DATE: CONCURRENTLY HERewith
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Sweeney, Patricia A.
REGISTRATION NUMBER: 32,733
REFERENCE/DOCKET NUMBER: 0578
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 248-4800
TELEFAX: (515) 248-4844
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1394 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-880-499-2
Query Match 94.0%; Score 1311; DB 3; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGATGCTCTATGAAAAAGATGATCAATGTCTATATCCGTTCTTAGGCTCC 60
DB 1 CCAATGGGTCTATGAAAAAGATGATCAATGTCTATATCCGTTCTTAGGCTCC 60
QY 61 CTCTCTGCTTATTAAGTATGATGATGATGATGATGATGATGATGATGATGAT 120
DB 61 CTCTCTGCTTATTAAGTATGATGATGATGATGATGATGATGATGATGATGAT 120
QY 121 CTCTGATGCTTATTAAGTATGATGATGATGATGATGATGATGATGATGAT 180
DB 121 CTCTGATGCTTATTAAGTATGATGATGATGATGATGATGATGATGATGAT 180
QY 181 CTGACCGAGGCGCATGACATGATGATGATGATGATGATGATGATGATGAT 240
DB 181 CTGACCGAGGCGCATGACATGATGATGATGATGATGATGATGATGATGAT 240
QY 241 CGAGAGCTATCGGCTGCTGATGATGATGATGATGATGATGATGATGATGAT 300
DB 241 CGAGAGCTATCGGCTGCTGATGATGATGATGATGATGATGATGATGATGAT 300
QY 301 TTAATATCTTATGATATTAATTAATTAATTAATTAATTAATTAATTAATTA 360
DB 301 TTAATATCTTATGATATTAATTAATTAATTAATTAATTAATTAATTAATTA 360
QY 361 GGGGCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
DB 361 GGGGCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
QY 421 TCCATATGAGCTATGATGATGATGATGATGATGATGATGATGATGATGAT 480
DB 421 TCCATATGAGCTATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 481 CGAAGCTATCATGATATGATGATGATGATGATGATGATGATGATGATGAT 540
DB 481 CGAAGCTATCATGATATGATGATGATGATGATGATGATGATGATGATGAT 540
QY 541 CAAGCATATTTGCAAAAGACAAATTAATTAATTAATTAATTAATTAATTA 600
DB 541 CAAGCATATTTGCAAAAGACAAATTAATTAATTAATTAATTAATTAATTA 600
QY 601 TTAATATGAGCTATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 601 TTAATATGAGCTATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 661 ATTGTTTTATATACATTTCTTCTTACATGATGATGATGATGATGATGAT 720
DB 661 ATTGTTTTATATACATTTCTTCTTACATGATGATGATGATGATGATGAT 720
QY 721 ATGACTATTAAGCTATTTATTAATTAAGACGATGATGATGATGATGATGAT 780
DB 721 ATGACTATTAAGCTATTTATTAATTAAGACGATGATGATGATGATGATGAT 780
QY 781 TTTCTGATTTTTTAAGAGCTATTTGGCAACCTGTTCTTCAAGAAATTTGAT 840
DB 781 TTTCTGATTTTTTAAGAGCTATTTGGCAACCTGTTCTTCAAGAAATTTGAT 840
QY 841 TTTCAAAAAAATAGTTATTTCTCTTTATTAATTAAGAAACATTTAGAAAT 900
DB 841 TTTCAAAAAAATAGTTATTTCTCTTTATTAATTAAGAAACATTTAGAAAT 900
QY 901 TGGCAATATGAGCTTATGATGATGATGATGATGATGATGATGATGATGAT 960
DB 901 TGGCAATATGAGCTTATGATGATGATGATGATGATGATGATGATGATGAT 960
QY 961 GCGAGCGGCTATTAATTAATTAAGCTGATGATGATGATGATGATGATGAT 1020
DB 961 GCGAGCGGCTATTAATTAATTAAGCTGATGATGATGATGATGATGATGAT 1020
QY 1021 TTTCTGATTTATTAATTAAGCTGATGATGATGATGATGATGATGATGAT 1080
DB 1021 TTTCTGATTTATTAATTAAGCTGATGATGATGATGATGATGATGATGAT 1080

QY 1081 TGAGATGTCGGTTTGGCAAGATAGCCACGTAATCATAGCTCATAGTGTCTAGTCA 1140
DB 1081 TGAGATGTCGGTTTGGCAAGATAGCCACGTAATCATAGCTCATAGTGTCTAGTCA 1140
QY 1141 GGTTCGGAGCTCTGCTGATCTCATCTCATGATGATGATGATGATGATGAT 1200
DB 1141 GGTTCGGAGCTCTGCTGATCTCATCTCATGATGATGATGATGATGATGAT 1200
QY 1201 TTGTTTCATGCTCCAAAGCTTGTCTATTTGAAACCAAGAGATTAATCTTCCAAACAT 1260
DB 1201 TTGTTTCATGCTCCAAAGCTTGTCTATTTGAAACCAAGAGATTAATCTTCCAAACAT 1260
QY 1261 CCATCTTACTCATGCACTTTCCATGCAACAGCAACATATGTTCTCTGAAAC 1311
DB 1261 CCATCTTACTCATGCACTTTCCATGCAACAGCAACATATGTTCTCTGAAAC 1311

RESULT 3
US-09-949-016-13563
; Sequence 13563, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13563
; LENGTH: 50263
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13563

Query Match 1.5%; Score 21; DB 4; Length 50263;
Best Local Similarity 100.0%; Pred. No. 4.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 861 TTTCTCTTATTAATTAAGAA 881
DB 40441 TTTCTCTTATTAATTAAGAA 40461

RESULT 4
US-09-949-016-124481/c
; Sequence 124481, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 124481
; LENGTH: 601
; TYPE: DNA

```

; ORGANISM: Human
US-09-949-016-124481

Query Match
Best Local Similarity 1.4%; Score 20; DB 4; Length 601;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 630 GATTTTCTTTTCATCTT 649
DB 363 GATTTTCTTTTCATCTT 344

RESULT 5
US-09-949-016-147075/c
; Sequence 147075, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 147075
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-147075

Query Match
Best Local Similarity 1.4%; Score 20; DB 4; Length 601;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 817 TTCTTCAAGATTTTGA 836
DB 353 TTCTTCAAGATTTTGA 334

RESULT 6
US-08-410-784A-3/c
; Sequence 3, Application US/08410784A
; Patent No. 5912413
; GENERAL INFORMATION:
; APPLICANT: MYERS, ALAN M.
; APPLICANT: JAMES, MARTHA G.
; TITLE OF INVENTION: ISOLATION OF SUL, A STARCH DEBRANCHING
; TITLE OF INVENTION: ENZYME, THE PRODUCT OF THE MAIZE GENE
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Weingarten, Schurgin, Gagnebin and Hayes LLP
; STREET: Ten Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/410,784A
; FILING DATE: 24-MAR-1995
; CLASSIFICATION: 800
```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Heine, Ph.D., Holliday C
; REGISTRATION NUMBER: 34,346
; REFERENCE/DOCKET NUMBER: ISU-002XX
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-2290
; TELEFAX: 617-451-0313
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2523 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
US-08-410-784A-3

Query Match
Best Local Similarity 1.4%; Score 20; DB 2; Length 2523;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 849 AAATTAGTTATTTCCTT 868
DB 206 AAATTAGTTATTTCCTT 187

RESULT 7
US-09-949-016-15250
; Sequence 15250, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 15250
; LENGTH: 21679
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15250

Query Match
Best Local Similarity 1.4%; Score 20; DB 4; Length 21679;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 630 GATTTTCTTTTCATCTT 649
DB 4880 GATTTTCTTTTCATCTT 4899

RESULT 8
US-09-949-016-12426/c
; Sequence 12426, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
```

```

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12426
; LENGTH: 65424
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(65424)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12426

Query Match          1.4%; Score 20; DB 4; Length 65424;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1041 GGAGAGAGAGAGAGAGAAAT 1060
Db      50936 GGAGAGAGAGAGAGAGAAAT 50917

RESULT 9
US-09-949-016-15868
; Sequence 15868, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15868
; LENGTH: 374159
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15868

Query Match          1.4%; Score 20; DB 4; Length 374159;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      817 TTTCCTTCAAGATTTTGA 836
Db      265055 TTTCCTTCAAGATTTTGA 265074

RESULT 10
US-09-949-016-17933/C
; Sequence 17933, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17933
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17933

Query Match          1.4%; Score 19; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      449 TTTAAAAAATACCAAG 467
Db      361 TTTAAAAAATACCAAG 343

RESULT 11
US-09-949-016-17934/C
; Sequence 17934, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17934
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17934

Query Match          1.4%; Score 19; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      449 TTTAAAAAATACCAAG 467
Db      394 TTTAAAAAATACCAAG 376

RESULT 12
US-09-949-016-41823/C
; Sequence 41823, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
```

```
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41823
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-41823

Query Match          1.4%; Score 19; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      449 TTTAAATAATTAACCAAG 467
      |||||
Db      361 TTTAAATAATTAACCAAG 343

RESULT 13
US-09-949-016-41824/c
; Sequence 41824, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41824
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-41824

Query Match          1.4%; Score 19; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      449 TTTAAATAATTAACCAAG 467
      |||||
Db      394 TTTAAATAATTAACCAAG 376

RESULT 14
US-09-949-016-145344/c
; Sequence 145344, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
```

```
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 145344
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-145344

Query Match          1.4%; Score 19; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1038 AAGGAGAGAGAGAGAGAG 1056
      |||||
Db      561 AAGGAGAGAGAGAGAGAG 543

RESULT 15
US-09-949-016-145345/c
; Sequence 145345, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 145345
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-145345

Query Match          1.4%; Score 19; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1038 AAGGAGAGAGAGAGAGAG 1056
      |||||
Db      553 AAGGAGAGAGAGAGAGAG 535
```

Search completed: September 16, 2005, 08:14:15
Job time : 179.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2005, 03:01:45 ; Search time 1291.5 Seconds
(without alignments)
7191.347 Million cell updates/sec

Title: US-10-713-381-1

Sequence score: 1394
1 ccacggcgcctcctacgaaaa.....cttttcgcacaccacacg 1394

Scoring table: Oligo NUC
Gapop 60.0 , Gapext 60.0

Searched: 7389322 seqs, 333128559 residues

Word size : 0

Total number of hits satisfying chosen parameters: 14778644

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA.*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US10J_PUBCOMB.seq:*
22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:*
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1394	100.0	1394	20	US-10-713-381-1
2	1311	94.0	1394	20	US-10-713-381-2
3	157	11.3	255	20	US-10-713-381-9
4	111	8.0	158	20	US-10-713-381-3
5	70	5.0	1663	20	US-10-739-930-1857
6	50	3.6	50	20	US-10-713-381-5
7	40	2.9	40	20	US-10-713-381-6

C	8	30	2.2	30	20	US-10-713-381-4	Sequence 4, Appl1
C	9	23	1.6	849	9	US-09-923-844B-5	Sequence 5, Appl1
C	10	22	1.6	194945	17	US-10-085-117-355	Sequence 355, App
C	11	21	1.5	333	20	US-10-357-930-56524	Sequence 56524, A
C	12	21	1.5	356	20	US-10-425-115-183040	Sequence 183040, A
C	13	21	1.5	386	20	US-10-425-115-117183	Sequence 117183, A
C	14	21	1.5	459	9	US-09-864-761-11465	Sequence 11465, A
C	15	21	1.5	1203	20	US-10-425-115-51470	Sequence 51470, A
C	16	21	1.5	1637	18	US-10-425-114-7365	Sequence 7365, App
C	17	21	1.5	1929	18	US-10-424-559-21199	Sequence 21199, A
C	18	21	1.5	9265	15	US-10-311-455-2125	Sequence 2125, App
C	19	21	1.5	9265	18	US-10-240-454-49	Sequence 49, Appl
C	20	21	1.5	18817	15	US-10-240-485-47	Sequence 47, Appl
C	21	21	1.5	26493	11	US-09-984-428-240	Sequence 114007, A
C	22	20	1.4	255	20	US-10-425-115-114007	Sequence 114007, A
C	23	20	1.4	257	18	US-10-424-559-119819	Sequence 119819, A
C	24	20	1.4	260	20	US-10-357-930-60745	Sequence 60745, A
C	25	20	1.4	311	20	US-10-425-115-176795	Sequence 176795, A
C	26	20	1.4	605	13	US-10-027-632-139460	Sequence 139460, A
C	27	20	1.4	605	13	US-10-027-632-139461	Sequence 139461, A
C	28	20	1.4	605	17	US-10-027-632-139460	Sequence 139460, A
C	29	20	1.4	605	17	US-10-027-632-139461	Sequence 139461, A
C	30	20	1.4	680	20	US-10-363-345A-4725	Sequence 4725, App
C	31	20	1.4	680	20	US-10-363-345A-4725	Sequence 4725, App
C	32	20	1.4	680	21	US-10-363-483A-4725	Sequence 4725, App
C	33	20	1.4	680	21	US-10-363-483A-4726	Sequence 4726, App
C	34	20	1.4	680	21	US-10-363-483A-4726	Sequence 4726, App
C	35	20	1.4	973	20	US-10-425-115-34238	Sequence 34238, A
C	36	20	1.4	1261	20	US-10-425-115-63769	Sequence 63769, A
C	37	20	1.4	1518	20	US-10-425-115-134230	Sequence 134230, A
C	38	20	1.4	51198	13	US-10-087-132-1729	Sequence 35698, A
C	39	20	1.4	54016	21	US-10-741-600-17886	Sequence 17886, A
C	40	20	1.4	72069	21	US-10-741-600-17886	Sequence 17886, A
C	41	20	1.4	91871	13	US-10-087-132-235	Sequence 235, App
C	42	20	1.4	149897	20	US-10-723-860-1744	Sequence 1744, App
C	43	20	1.4	149897	22	US-10-755-149-675	Sequence 675, App
C	44	20	1.4	337344	19	US-10-388-838-58	Sequence 58, Appl
C	45	20	1.4	465237	9	US-09-933-267A-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-10-713-381-1
; Sequence 1, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-1

Query Match 100.0%; Score 1394; DB 20; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCATGCTGCTCTATGAAAAGATGATGATGCTCTATGCTCTTTCTTAGGCTCC 60

Db 1 CCAATGCTCTCTATGAAAAAGATGAGTACAAATGTGTCTATATCCGTTTCTTAGGCTCC 60
Qy 61 CTTCTTCTGCTTATTAATGACTGAATCGGGGTTTACAAAACTTCCAGGTGATGAT 120
Db 61 CTTCTTCTGCTTATTAATGACTGAATCGGGGTTTACAAAACTTCCAGGTGATGAT 120
Qy 121 CTCATGTTTCACTTCCACCTCGGGGTGACATTTCTTGATGTGGTGGTCCCAT 180
Db 121 CTCATGTTTCACTTCCACCTCGGGGTGACATTTCTTGATGTGGTGGTCCCAT 180
Qy 181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGCTTTCGATGCCCCA 240
Db 181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGCTTTCGATGCCCCA 240
Qy 241 CGAGAGTATCGGGTGTGTGATCCAGGGGATATATGTCCCAATGCTCACTTAA 300
Db 241 CGAGAGTATCGGGTGTGTGATCCAGGGGATATATGTCCCAATGCTCACTTAA 300
Qy 301 TTATATCTTGTAGATATTAATTTTGAATAAACAATTAATCTTGTGTGA 360
Db 301 TTATATCTTGTAGATATTAATTTTGAATAAACAATTAATCTTGTGTGA 360
Qy 361 GGGCTCAGCATAGATTTTGGCTTAGGGCCAGAAAATGCGAGGACGACCATGTCTAGTG 420
Db 361 GGGCTCAGCATAGATTTTGGCTTAGGGCCAGAAAATGCGAGGACGACCATGTCTAGTG 420
Qy 421 TCCATCTATTGGCATCCCAAGAAAGATTTAAACCAATTAACCAAGTAACTATCACT 480
Db 421 TCCATCTATTGGCATCCCAAGAAAGATTTAAACCAATTAACCAAGTAACTATCACT 480
Qy 481 CGAAGCTATCATGTATGTTTAAAGAAACATCTATTAACCAAGATCCTTTAAATAA 540
Db 481 CGAAGCTATCATGTATGTTTAAAGAAACATCTATTAACCAAGATCCTTTAAATAA 540
Qy 541 CAAGCATATTTGCAAAAGAGCAAAATTAATGTTACAGTTTAAACCAATCTAAGACGACAA 600
Db 541 CAAGCATATTTGCAAAAGAGCAAAATTAATGTTACAGTTTAAACCAATCTAAGACGACAA 600
Qy 601 TTATATCGAAGGTAGCTATGACGTTCCAGATTTTCTTTTCACTCTGTATTTTGT 660
Db 601 TTATATCGAAGGTAGCTATGACGTTCCAGATTTTCTTTTCACTCTGTATTTTGT 660
Qy 661 ATTGTTTTATATACATTTTCTCTCTTCAATAGAGTATTTCTTCCGATTTTAA 720
Db 661 ATTGTTTTATATACATTTTCTCTCTTCAATAGAGTATTTCTTCCGATTTTAA 720
Qy 721 ATGACTATAAAGTCATTTTATATTAATAAGACGACATGTGTAGATTCTCGTTCAAAAATC 780
Db 721 ATGACTATAAAGTCATTTTATATTAATAAGACGACATGTGTAGATTCTCGTTCAAAAATC 780
Qy 781 TTTCTGATTTTAAAGCTAGTTTGGCAACCTGTCTTCAAGAAATTTGATTTT 840
Db 781 TTTCTGATTTTAAAGCTAGTTTGGCAACCTGTCTTCAAGAAATTTGATTTT 840
Qy 841 TTCAAAAAAATAGTTTATTTCTCTTATTAATAAGAAACACTTAGAAAAATAGAGT 900
Db 841 TTCAAAAAAATAGTTTATTTCTCTTATTAATAAGAAACACTTAGAAAAATAGAGT 900
Qy 901 TGGCAGACTAGCCCTAGAAATGTTTTCCCAATAAATTAACAATCACTGTATTAATTTT 960
Db 901 TGGCAGACTAGCCCTAGAAATGTTTTCCCAATAAATTAACAATCACTGTATTAATTTT 960
Qy 961 GCGAGGCCCATTAATTAATTAACGAAATCTGAAATCGAGCCAAACCAATCTGACTAT 1020
Db 961 GCGAGGCCCATTAATTAATTAACGAAATCTGAAATCGAGCCAAACCAATCTGACTAT 1020
Qy 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAAATAGATTAAAGTCTTGTCCC 1080
Db 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAAATAGATTAAAGTCTTGTCCC 1080
Qy 1081 TGAGATGTGCGGTTTGGCAACGATAGCCCGTAATCATAGTGTGCTAGCTGA 1140
Db 1081 TGAGATGTGCGGTTTGGCAACGATAGCCCGTAATCATAGTGTGCTAGCTGA 1140

Db 1081 TGAGATGTGCGGTTTGGCAACGATAGCCCGTAATCATAGTGTGCTAGCTGA 1140
Qy 1141 GGTGGGACGCTCTGTGTCTATCTACATGAGATACATGCTTGTAAACCGTTGCTC 1200
Db 1141 GGTGGGACGCTCTGTGTCTATCTACATGAGATACATGCTTGTAAACCGTTGCTC 1200
Qy 1201 TTGTTTCATGCTGCAAGCCCTGCTATTTCTGAAACAAAGAGADACTACTCCCAACAT 1260
Db 1201 TTGTTTCATGCTGCAAGCCCTGCTATTTCTGAAACAAAGAGADACTACTCCCAACAT 1260
Qy 1261 CCATCTTACTCATGCAACTTCCATGCAACACGCAATATGTTTCTGAAACCAATTCAT 1320
Db 1261 CCATCTTACTCATGCAACTTCCATGCAACACGCAATATGTTTCTGAAACCAATTCAT 1320
Qy 1321 AAAAGATCAACAGCTAGGTTCTCCGCTACTTCCCTCTCTCTGCGGATCTTTT 1380
Db 1321 AAAAGATCAACAGCTAGGTTCTCCGCTACTTCCCTCTCTCTGCGGATCTTTT 1380
Qy 1381 CGTCCACGACATG 1394
Db 1381 CGTCCACGACATG 1394

RESULT 2
US-10-713-381-2
; Sequence 2, Application US/10713381
; Publication No. US2004022131A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFEMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: USING SAME
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-2

Query Match 94.0%; Score 1311; DB 20; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCAATGCTCTCTATGAAAAAGATGAGTACAAATGTGTCTATATCCGTTTCTTAGGCTCC 60
Db 1 CCAATGCTCTCTATGAAAAAGATGAGTACAAATGTGTCTATATCCGTTTCTTAGGCTCC 60
Qy 61 CTTCTTCTGCTTATTAATGACTGAATCGGGGTTTACAAAACTTCCAGGTGATGAT 120
Db 61 CTTCTTCTGCTTATTAATGACTGAATCGGGGTTTACAAAACTTCCAGGTGATGAT 120
Qy 121 CTCATGTTTCACTTCCACCTCGGGGTGACATTTCTTGATGTGGTGGTCCCAT 180
Db 121 CTCATGTTTCACTTCCACCTCGGGGTGACATTTCTTGATGTGGTGGTCCCAT 180
Qy 181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGCTTTCGATGCCCCA 240
Db 181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGCTTTCGATGCCCCA 240
Qy 241 CGAGAGTATCGGGTGTGTGATCCAGGGGATATATGTCCCAATGCTCACTTAA 300
Db 241 CGAGAGTATCGGGTGTGTGATCCAGGGGATATATGTCCCAATGCTCACTTAA 300
Qy 301 TTATATCTTGTAGATATTAATTTTGAATAAACAATTAATCTTGTGTGA 360
Db 301 TTATATCTTGTAGATATTAATTTTGAATAAACAATTAATCTTGTGTGA 360

```
Db      301 TTAATTCTTTAGATATTATTTTGGAAAAATTAACAATTAATCTTATCTTTGGTA 360
Qy      361 GGGCCCTCAGCATGATTTTGGCTTAGGGCCAGAAAATGCGAGAGCAGCCATGCTAGTG 420
Db      361 GGGCCCTCAGCATGATTTTGGCTTAGGGCCAGAAAATGCGAGAGCAGCCATGCTAGTG 420
Qy      421 TCCACTATTGGCATACCCAGAAACAAGATTAAAAAATAACAAAGTAATCATCCACT 480
Db      421 TCCACTATTGGCATACCCAGAAACAAGATTAAAAAATAACAAAGTAATCATCCACT 480
Qy      481 CGAAAGCTATCATGATATGTTTAAAGAAACATGATTTAAACACGATCCCTTAAAAAA 540
Db      481 CGAAAGCTATCATGATATGTTTAAAGAAACATGATTTAAACACGATCCCTTAAAAAA 540
Qy      541 CAAGCATATTTGCAAAAGACAAATATATGTACAGTTTACAAACATCTAAGACGACAA 600
Db      541 CAAGCATATTTGCAAAAGACAAATATATGTACAGTTTACAAACATCTAAGACGACAA 600
Qy      601 TTTATATCGAAAGGTAAAGCTATGACGTTCAAGATTTTCTTTTCAATCTTGTATTTGTT 660
Db      601 TTTATATCGAAAGGTAAAGCTATGACGTTCAAGATTTTCTTTTCAATCTTGTATTTGTT 660
Qy      661 ATGTGTTTTATATACATTTTCTTCTTACAAATAGATGATTTTCTTCCGATTTTAA 720
Db      661 ATGTGTTTTATATACATTTTCTTCTTACAAATAGATGATTTTCTTCCGATTTTAA 720
Qy      721 ATGACTATTAAGTCAATTTTATATTAAGAGACGACATGCTAGATTTCTGTTCAAAAATC 780
Db      721 ATGACTATTAAGTCAATTTTATATTAAGAGACGACATGCTAGATTTCTGTTCAAAAATC 780
Qy      781 TTTCTGATTTTTTAAAGAGTACTAGTTGGCAACCTGTTTCTTCAAAAGATTTGATTTT 840
Db      781 TTTCTGATTTTTTAAAGAGTACTAGTTGGCAACCTGTTTCTTCAAAAGATTTGATTTT 840
Qy      841 TTTCAAAAAAATAGTTTATTTCTCTTATTAATAAGAAAAACATAGAAAAATAGAGT 900
Db      841 TTTCAAAAAAATAGTTTATTTCTCTTATTAATAAGAAAAACATAGAAAAATAGAGT 900
Qy      901 TGGCAGACTAGCCCTAGAAATGTTTTCCCAATAAATTAACAATCACTGATTAATTAATTTG 960
Db      901 TGGCAGACTAGCCCTAGAAATGTTTTCCCAATAAATTAACAATCACTGATTAATTAATTTG 960
Qy      961 GCCAGCCCCCAATAATTTTAAACCGAAACTGAATCGAGCGAAACCAAAATCTGACCTAT 1020
Db      961 GCCAGCCCCCAATAATTTTAAACCGAAACTGAATCGAGCGAAACCAAAATCTGACCTAT 1020
Qy      1021 TTCTCTAGATTAGTAAAGAGGAGAGAGAGAGAAATCACTTTTAAGTCAATGTCCTC 1080
Db      1021 TTCTCTAGATTAGTAAAGAGGAGAGAGAGAGAAATCACTTTTAAGTCAATGTCCTC 1080
Qy      1081 TGAGATGTGGGTTTGGCAACGATAGCAACCGTATCATAGCTCATAGGTGCTTACGTCA 1140
Db      1081 TGAGATGTGGGTTTGGCAACGATAGCAACCGTATCATAGGTGCTTACGTCA 1140
Qy      1141 GGTTCGGCAGCTCTGCTGTCATCTCAATGCGATACATGCTGCTGTTCAACCGTTCGTC 1200
Db      1141 GGTTCGGCAGCTCTGCTGTCATCTCAATGCGATACATGCTGCTGTTCAACCGTTCGTC 1200
Qy      1201 TTGTTCCATCGTCCAGAGCTTGGCTATTCTGAACCAAGAGATACCTATCCCAAAACAT 1260
Db      1201 TTGTTCCATCGTCCAGAGCTTGGCTATTCTGAACCAAGAGATACCTATCCCAAAACAT 1260
Qy      1261 CCATCTTACTCATGCAACTTCCATGCAACCGCACATATGTTTCTGTAAC 1311
Db      1261 CCATCTTACTCATGCAACTTCCATGCAACCGCACATATGTTTCTGTAAC 1311
```

RESULT 3
US-10-713-381-9
; Sequence 9, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:

```
; APPLICANT: ALBERTSEN, MARC C.  
; APPLICANT: FOX, TIMOTHY W.  
; APPLICANT: GARNAT, CARL W.  
; APPLICANT: HUFFMAN, GARY  
; APPLICANT: KENDALL, TIMMY L.  
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF  
; FILE REFERENCE: 578R  
; CURRENT APPLICATION NUMBER: US/10/713,381  
; PRIOR FILING DATE: 2003-11-14  
; PRIOR APPLICATION NUMBER: 08/880,499  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 9  
; LENGTH: 255  
; TYPE: DNA  
; ORGANISM: Zea mays  
US-10-713-381-9  
  
Query Match 11.3%; Score 157; DB 20; Length 255;  
Best Local Similarity 100.0%; Pred. No. 4.8e-68;  
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy      1155 CGTGTCACTCATGACATGCGATTAATCACTGTTTCAACCGTTCGTTGTCATGTC 1214  
Db      15 CGTGTCACTCATGACATGCGATTAATCACTGTTTCAACCGTTCGTTGTCATGTC 74  
  
Qy      1215 AAGCTTGCTTATTTGGAACCAAGAGATACCTATCCCAACATTCATCTTACTCATG 1274  
Db      75 AAGCTTGCTTATTTGGAACCAAGAGATACCTATCCCAACATTCATCTTACTCATG 134  
  
Qy      1275 CAATTCATGCAACACGACATATGTTTCTGTAAC 1311  
Db      135 CAATTCATGCAACACGACATATGTTTCTGTAAC 171  
  
RESULT 4  
US-10-713-381-3  
; Sequence 3, Application US/10713381  
; Publication No. US20040221331A1  
; GENERAL INFORMATION:  
; APPLICANT: ALBERTSEN, MARC C.  
; APPLICANT: FOX, TIMOTHY W.  
; APPLICANT: GARNAT, CARL W.  
; APPLICANT: HUFFMAN, GARY  
; APPLICANT: KENDALL, TIMMY L.  
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF  
; FILE REFERENCE: 578R  
; CURRENT APPLICATION NUMBER: US/10/713,381  
; PRIOR FILING DATE: 2003-11-14  
; PRIOR APPLICATION NUMBER: 08/880,499  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 158  
; TYPE: DNA  
; ORGANISM: Zea mays  
US-10-713-381-3  
  
Query Match 8.0%; Score 111; DB 20; Length 158;  
Best Local Similarity 100.0%; Pred. No. 6.9e-45;  
Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy      1201 TTGTTCCATCGTCCAGAGCTTGGCTATTCTGAACCAAGAGATTAATCTCCCAAAACAT 1260  
Db      48 TTGTTCCATCGTCCAGAGCTTGGCTATTCTGAACCAAGAGATTAATCTCCCAAAACAT 107  
  
Qy      1261 CCATCTTACTCATGCAACTTCCATGCAACCGCACATATGTTTCTGTAAC 1311  
Db      108 CCATCTTACTCATGCAACTTCCATGCAACCGCACATATGTTTCTGTAAC 158
```

```
RESULT 5
US-10-739-930-1857
; Sequence 1857, Application US/10739930
; Publication No. US20040216150A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; PRIOR FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 1857
; LENGTH: 1663
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: ZEAWA-23APR03-CLUSTER1429_8
US-10-739-930-1857

Query Match          5.0%; Score 70; DB 20; Length 1663;
Best Local Similarity 100.0%; Pred. No. 3.8e-24;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1313 AATCCATTAAAGATCAACAGCTAGCGTCTCCGCTGCTCTCTCTGCGG 1372
DB      1 AATCCATTAAAGATCAACAGCTAGCGTCTCCGCTGCTCTCTCTGCGG 60

QY      1373 ATCTTTTCG 1382
DB      61 ATCTTTTCG 70

RESULT 6
US-10-713-381-5
; Sequence 5, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-5

Query Match          3.6%; Score 50; DB 20; Length 50;
Best Local Similarity 100.0%; Pred. No. 3.4e-14;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1229 CTGAACCAAGAGATACCTTACCTCCAAACATTCATCTTACTGCAAC 1278
DB      1 CTGAACCAAGAGATACCTTACCTCCAAACATTCATCTTACTGCAAC 50

RESULT 7
US-10-713-381-6
; Sequence 6, Application US/10713381
; Publication No. US20040221331A1
```

```
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-6

Query Match          2.9%; Score 40; DB 20; Length 40;
Best Local Similarity 100.0%; Pred. No. 3.6e-09;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1239 AGATACCTACTCCGCAACATTCATCTTACTGCAAC 1278
DB      1 AGATACCTACTCCGCAACATTCATCTTACTGCAAC 40

RESULT 8
US-10-713-381-4
; Sequence 4, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-4

Query Match          2.2%; Score 30; DB 20; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.00039;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1179 CATGCTGTTCACCGTTCGTCTGTTCCTCA 1208
DB      1 CATGCTGTTCACCGTTCGTCTGTTCCTCA 30

RESULT 9
US-09-923-844B-5/c
; Sequence 5, Application US/09923844B
; Patent No. US2002016143A1
; GENERAL INFORMATION:
; APPLICANT: Pioneer Hi-Bred International, Inc.
; APPLICANT: Bao, Zhongmeng
; APPLICANT: Lu, Guihua
; TITLE OF INVENTION: Sclerotinia-inducible Genes and
```

```
/ TITLE OF INVENTION: Promoters and Their Uses
/ FILE REFERENCE: 35718/234631
/ CURRENT APPLICATION NUMBER: US/09/923, 844B
/ PRIOR FILING DATE: 2001-08-07
/ PRIOR APPLICATION NUMBER: US 60/224,603
/ PRIOR FILING DATE: 2000-08-11
/ NUMBER OF SEQ ID NOS: 20
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 5
/ LENGTH: 849
/ TYPE: DNA
/ ORGANISM: Helianthus annuus
US-09-923-844B-5

Query Match          1.6%; Score 23; DB 9; Length 849;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      854 AGTTATTTCTCTTATATAAT 876
DB      556 AGTTATTTCTCTTATATAAT 534

RESULT 10
US-10-085-117-355/c
/ Sequence 355, Application US/10085117
/ Publication No. US2003023334A1
/ GENERAL INFORMATION:
/ APPLICANT: Morris, David W.
/ APPLICANT: Engelhard, Eric K.
/ TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
/ FILE REFERENCE: 529452000121
/ CURRENT APPLICATION NUMBER: US/10/085,117
/ PRIOR FILING DATE: 2002-02-27
/ PRIOR APPLICATION NUMBER: US 09/798,586
/ PRIOR FILING DATE: 2001-03-02
/ NUMBER OF SEQ ID NOS: 361
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 355
/ LENGTH: 194945
/ TYPE: DNA
/ ORGANISM: Mus musculus
/ FEATURE:
/ NAME/KEY: variation
/ LOCATION: (1)...(194945)
/ OTHER INFORMATION: n = any nucleotide
US-10-085-117-355

Query Match          1.6%; Score 22; DB 17; Length 194945;
Best Local Similarity 100.0%; Pred. No. 8.4;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      113 TGCATGATCTCCATGTTCCACT 134
DB      28972 TGCATGATCTCCATGTTCCACT 28951

RESULT 11
US-10-357-930-56524
/ Sequence 56524, Application US/10357930
/ Publication No. US20040259086A1
/ GENERAL INFORMATION:
/ APPLICANT: Schlegel, Robert
/ APPLICANT: Endege, Wilson
/ APPLICANT: Monahan, John
/ TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
/ TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
/ TITLE OF INVENTION: HUMAN PROSTATE CANCER
/ FILE REFERENCE: MRI-007BCN
/ CURRENT APPLICATION NUMBER: US/10/357,930
/ PRIOR FILING DATE: 2003-02-04
/ PRIOR APPLICATION NUMBER: 09/785,276
/ PRIOR FILING DATE: 2003-02-16
```

```
/ PRIOR APPLICATION NUMBER: 60/183,319
/ PRIOR FILING DATE: 2000-02-17
/ PRIOR APPLICATION NUMBER: 60/189,862
/ PRIOR FILING DATE: 2000-03-16
/ PRIOR APPLICATION NUMBER: 60/207,454
/ PRIOR FILING DATE: 2000-05-25
/ PRIOR APPLICATION NUMBER: 60/211,314
/ PRIOR FILING DATE: 2000-06-09
/ PRIOR APPLICATION NUMBER: 60/219,007
/ PRIOR FILING DATE: 2000-07-18
/ PRIOR APPLICATION NUMBER: 60/255,281
/ PRIOR FILING DATE: 2000-12-13
/ NUMBER OF SEQ ID NOS: 62232
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 56524
/ LENGTH: 333
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-357-930-56524

Query Match          1.5%; Score 21; DB 20; Length 333;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      93 TTACAAAAAATCTCCACGGGT 113
DB      89 TTACAAAAAATCTCCACGGGT 109

RESULT 12
US-10-425-115-183040
/ Sequence 183040, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(53222)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ PRIOR FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 183040
/ LENGTH: 356
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MRT4577_98517C.1
US-10-425-115-183040

Query Match          1.5%; Score 21; DB 20; Length 356;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      632 TTTTCTTTTCATCTTGTT 652
DB      72 TTTTCTTTTCATCTTGTT 92

RESULT 13
US-10-425-115-117183/c
/ Sequence 117183, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants
/ FILE REFERENCE: 38-21(53222)B
```

```
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 117183
/ LENGTH: 386
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MRT4577_3835C.1
US-10-425-115-117183

Query Match          1.5%; Score 21; DB 20; Length 386;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 371 ATGATTTTCGCTTAGGCC 391
DB 367 ATGATTTTCGCTTAGGCC 347

RESULT 14
US-09-864-761-11465/c
/ Sequence 11465, Application US/09864761
/ Patent No. US20020048763A1
/ GENERAL INFORMATION:
/ APPLICANT: Penn, Sharron G.
/ APPLICANT: Rank, David R.
/ APPLICANT: Hanzel, David K.
/ APPLICANT: Chen, Wenhang
/ TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
/ FILE REFERENCE: Aecomica-X-1
/ CURRENT APPLICATION NUMBER: US/09/864,761
/ PRIOR APPLICATION NUMBER: US 60/180,312
/ PRIOR FILING DATE: 2000-02-04
/ PRIOR APPLICATION NUMBER: US 60/207,456
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: US 09/632,366
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR APPLICATION NUMBER: GB 24263.6
/ PRIOR FILING DATE: 2000-10-04
/ PRIOR APPLICATION NUMBER: US 60/236,359
/ PRIOR FILING DATE: 2000-09-27
/ PRIOR APPLICATION NUMBER: PCT/US01/00664
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00666
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00667
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00664
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00668
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00663
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00662
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00661
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00670
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: US 60/234,687
/ PRIOR FILING DATE: 2000-09-21
/ PRIOR APPLICATION NUMBER: US 09/608,408
/ PRIOR FILING DATE: 2000-06-30
/ PRIOR APPLICATION NUMBER: US 09/774,203
/ PRIOR FILING DATE: 2001-01-29
/ NUMBER OF SEQ ID NOS: 49117
/ SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
/ SEQ ID NO 11465
```

```
/ LENGTH: 459
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO AC000100.2
/ OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.59
/ OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.59
/ OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.59
/ OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.6
/ OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.64
/ OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.2
US-09-864-761-11465

Query Match          1.5%; Score 21; DB 9; Length 459;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1037 AAAGGAGAGAGAGAGAGA 1057
DB 118 AAAGGAGAGAGAGAGAGA 98

RESULT 15
US-10-425-115-51470/c
/ Sequence 51470, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovacic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(53222)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 51470
/ LENGTH: 1203
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (1)..(1203)
/ OTHER INFORMATION: unsure at all n locations
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MRT4577_146936C.1
US-10-425-115-51470

Query Match          1.5%; Score 21; DB 20; Length 1203;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 848 AAAATTAGTTATTTCTCTT 868
DB 1129 AAAATTAGTTATTTCTCTT 1109

Search completed: September 16, 2005, 12:49:09
Job time : 1294.5 secs
```



```
QY 464 AAAGTAACATCACTCGAAGGATCATGTAATGTTTAAAGAAATCTATTAAACC 523
DB 61 AAAGTAACATCACTCGAAGGATCATGTAATGTTTAAAGAAATCTATTAAACC 120
QY 524 AGCATCTCTTAAACAAAGCATATTTGCAAAAGACAAATATGTTACAGTTTACAA 583
DB 121 AGCATCTCTTAAACAAAGCATATTTGCAAAAGACAAATATGTTACAGTTTACAA 180
QY 584 CATCTAAGACGCAAAATATATCGAAAGGATGATGACGTTTCAGATTTTCTTTTC 643
DB 181 CATCTAAGACGCAAAATATATCGAAAGGATGATGACGTTTCAGATTTTCTTTTC 240
QY 644 ATTCTGTATTTTGTATGTTTATATACATTTTCTCTCTCAATAGAGTATTT 703
DB 241 ATTCTGTATTTTGTATGTTTATATACATTTTCTCTCTCAATAGAGTATTT 300
QY 704 TCTTCGATTTTATTAATGATATTAAGTCAATTTTATTAAGACGACGATGCTAG 763
DB 301 TCTTCGATTTTATTAATGATATTAAGTCAATTTTATTAAGACGACGATGCTAG 360
QY 764 ATTCTGTATTTTGTATGTTTATTAAGAGTATTTGCAACCTGTTCTTT 823
DB 361 ATTCTGTATTTTGTATGTTTATTAAGAGTATTTGCAACCTGTTCTTT 420
QY 824 CAAGAATTTTGAATTTTCAAAAAAATAGTTTATTTCTCTTATTAATAAGAAAC 883
DB 421 CAAGAATTTTGAATTTTCAAAAAAATAGTTTATTTCTCTTATTAATAAGAAAC 480
QY 884 ACTTGAATAATAGAGTGGCAAGCTAGCCCTAGAAATGTTTCCCAATAATTAACATCA 943
DB 481 ACTTGAATAATAGAGTGGCAAGCTAGCCCTAGAAATGTTTCCCAATAATTAACATCA 540
QY 944 CTGTGTATATTTTGGCCAGCCCAATAATTAATTAAGCAATGTAATCGAGCA 1003
DB 541 CTGTGTATATTTTGGCCAGCCCAATAATTAATTAAGCAATGTAATCGAGCA 600
QY 1004 AACCAATCTGAGCTATTTCTCTAGATTAATAAGAGAGAGAGAGAGAGAGAGAGAG 1063
DB 601 AACCAATCTGAGCTATTTCTCTAGATTAATAAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 1064 TTTTAAGTCAATGTCCTGAGATGTCGGTGTGGCAAGATGCAACCGTATCATAGCT 1123
DB 661 TTTTAAGTCAATGTCCTGAGATGTCGGTGTGGCAAGATGCAACCGTATCATAGCT 720
QY 1124 CATAGGTCCTAGGTCAGGTCGGAGCTCTCGGTCATCTCAATGAGCACTACTACATGC 1183
DB 721 CATAGGTCCTAGGTCAGGTCGGAGCTCTCGGTCATCTCAATGAGCACTACTACATGC 780
QY 1184 TTGTTCACCGTTCGTTCGTTCATGTCGTCAGACCTTGCCATTCGTAACCAAGAGAT 1243
DB 781 TTGTTCACCGTTCGTTCGTTCATGTCGTCAGACCTTGCCATTCGTAACCAAGAGAT 840
QY 1244 ACTTACTCCCAAAATCTCTTACTCATGCAACTTCATGCAAAACGCAATATGTT 1303
DB 841 ACTTACTCCCAAAATCTCTTACTCATGCAACTTCATGCAAAACGCAATATGTT 900
QY 1304 TCCTGAACCA 1313
DB 901 TCCTGAACCA 910

RESULT 2
CG224225 915 bp DNA linear GSS 22-AUG-2003
LOCUS OGIAG081Y_ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMA0716B15,
DEFINITION genomic survey sequence.
ACCESSION CG224225
VERSION CG224225.1 GI:34124113
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```

```
REFERENCE
1 (bases 1 to 915)
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
AUTHORS
Whiteley,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Numborg,A., Robbins,D. and Lakey,N.
CONSORTIUM for Maize Genomics
JOURNAL
Unpublished (2002)
COMMENT
Other GSSs: OGIAG08TH
Contact: Cathy Whiteley
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteley@cigr.org
Seq primer: 7F
Classes: sheared ends.
FEATURES
source
location/Qualifiers
1..915
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0716B15"
/clone_1ib="ZM_0.7_1.5_KB"
/note="Vector: pBCK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"
ORIGIN
Query Match 53.4%; Score 745; DB 9; Length 915;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 745; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 645 TTCTGTATTTTGTATGTTTATATATACATTTCTCTCTCAATAGAGTATTT 704
DB 24 TTCTGTATTTTGTATGTTTATATATACATTTCTCTCTCAATAGAGTATTT 83
QY 705 CTTCGATTTTATTAATGATATTAAGTATTTTATTAAGACAGCATGCTAGAGA 764
DB 84 CTTCGATTTTATTAATGATATTAAGTATTTTATTAAGACAGCATGCTAGAGA 143
QY 765 TTCTGTCATAAAATCTTTCTGATTTTAAAGACTAGTTGGCAACCTGTTCTTTC 824
DB 144 TTCTGTCATAAAATCTTTCTGATTTTAAAGACTAGTTGGCAACCTGTTCTTTC 203
QY 825 AAAGAATTTGATTTTCAAAAAAATAGTTTATTTCTCTTATTAATAAGAAACA 884
DB 204 AAAGAATTTGATTTTCAAAAAAATAGTTTATTTCTCTTATTAATAAGAAACA 263
QY 885 CTTAGAAAATAGAGTGGCAGACTAGCCCTAGAAATGTTTCCCAATAATTAACATCAC 944
DB 264 CTTAGAAAATAGAGTGGCAGACTAGCCCTAGAAATGTTTCCCAATAATTAACATCAC 323
QY 945 TTGTATTAATTTTGGCCAGCCCAATAATTAATTTAAACGAAATCGAGCGAA 1004
DB 324 TTGTATTAATTTTGGCCAGCCCAATAATTAATTTAAACGAAATCGAGCGAA 383
QY 1005 ACCAATCTGAGCTATTTCTCTAGATTTAGTAAAGAGAGAGAGAGAGAGAGATTCGT 1064
DB 384 ACCAATCTGAGCTATTTCTCTAGATTTAGTAAAGAGAGAGAGAGAGAGATTCGT 443
QY 1065 TTTAAGTCAATGTCCTGAGATGTCGGTGTGGCAAGATGCAACCGTATCATAGCTC 1124
DB 444 TTTAAGTCAATGTCCTGAGATGTCGGTGTGGCAAGATGCAACCGTATCATAGCTC 503
QY 1125 ATAGTGCTTACGTCAGGTCGGCAGCTCTCTGTTCATCTCAATGAGCATATCATGCT 1184
DB 504 ATAGTGCTTACGTCAGGTCGGCAGCTCTCTGTTCATCTCAATGAGCATATCATGCT 563
QY 1185 TGTTCAACCGTTCGTCTGTCATGTCGTCAGACCTTGCTTTCGAAACCAAGAGGATA 1244
DB 564 TGTTCAACCGTTCGTCTGTCATGTCGTCAGACCTTGCTTTCGAAACCAAGAGGATA 623
```


Oy	1245	CCTACTGCCAAACATTCATCTTACTCATGTGCACTTCCATGCAACAGACACATATGTT	1304
Db	624	CCTACTGCCAAACATTCATCTTACTCATGTGCACTTCCATGCAACAGACACATATGTT	683
Oy	1305	CCTGAACCAATTCATTTAAAGATCAGACAGCTAGCGTTCTCCCGTAGCTTCCCTCTC	1364
Db	684	CCTGAACCAATTCATTTAAAGATCAGACAGCTAGCGTTCTCCCGTAGCTTCCCTCTC	743
Oy	1365	CTGTGCCGATCTTTTTCGTCCACCA	1389
Db	744	CTGTGCCGATCTTTTTCGTCCACCA	768
RESULT 3			
LOCUS	CC656939	687 bp	DNA
DEFINITION	OGMDQ20TV ZM 0.7.1.5_KB Zea mays genomic clone ZM8BMA0554D15,	linear	GSS 19-JUN-2003
ACCESSION	CC656939	genomic survey sequence.	
VERSION	CC656939.1	GI:32060231	
KEYWORDS	GSS.		
SOURCE	Zea mays		
ORGANISM	Zea mays		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoidae; Andropogoneae; Zea.		
AUTHORS	Whiteclaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Renick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nuneberg,A., Robbins,D. and Lakey,N.		
TITLE	Consortium for Maize Genomics		
JOURNAL	Unpublished (2002)		
COMMENT	Other GSSs: OGMDQ20TM		
	Contact: Cathy Whiteclaw		
	TIGR		
	9712 Medical Center Drive, Rockville, MD 20850, USA		
	Tel: 301-838-5843		
	Fax: 301-838-0208		
	Email: whiteclaw@tigr.org		
	Seq primer: TP		
	Class: sheared ends.		
FEATURES	Location/Qualifiers		
source	1..687		
	/organism="Zea mays"		
	/mol_type="genomic DNA"		
	/strain="B73"		
	/db_xref="taxon:4577"		
	/clone="ZM8BMA0554D15"		
	/clone_11b="ZM 0.7.1.5_KB"		
	/note="vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb		
	methylation filtered genomic DNA library"		
ORIGIN			
Query Match	34.9%	Score 487;	DB 9; Length 687;
Best Local Similarity	100.0%;	Pred. No. 1.2e-217;	
Matches 487; Conservative	0; Mismatches	0; Indels	0; Gaps 0;
Oy	903	CCAACTGCGCCTAGATAGTGTTCCTCAATAATTACAATCATCTGTATATTTATTTGGC	962
Db	673	CCAGACTAGCCCTAGAAATGTTTCCCAATTAATTACAATCATCTGTATATTTATTTGGC	614
Oy	963	CAGCCCCATAATTATTTAAACCGAACTGAATTCGAGCGAAACCAATCTGAGCTATTT	1022
Db	613	CAGCCCCATAATTATTTAAACCGAACTGAATTCGAGCGAAACCAATCTGAGCTATTT	554
Oy	1023	CTTAGATTAGTAAAAAGGAGAGAGAGAGAGAAAGAAATCACTTTTAAGTCATTTGTCCTG	1082
Db	553	CTTAGATTAGTAAAAAGGAGAGAGAGAGAGAAATCACTTTTAAGTCATTTGTCCTG	494
Oy	1083	AGATGTGGGTTTGGCAAGATAGCAACCGTAATCATATGATCATAGTGCTAGCTAGG	1142
Db	493	AGATGTGGGTTTGGCAAGATAGCAACCGTAATCATATGATCATAGTGCTAGCTAGG	434

Qy	1143	TTGGGAGGCTCGTGTGCATCTGCACATGGGATACATGCTTGTGAACGGTTCGCTT	1202
Db	433	TTGGGAGGCTCGTGTGCATCTGCACATGGGATACATGCTTGTGAACGGTTCGCTT	374
Qy	1203	GTTCATCTGTCCAGCGCTTGCTATTCTGAAACCAAGAGATACCTACTCCCAACAAATCC	1262
Db	373	GTTCATCTGTCCAGCGCTTGCTATTCTGAAACCAAGAGATACCTACTCCCAACAAATCC	314
Qy	1263	ATCTTACTCATGCACTTCCATGCAAAACAGCACATATGTTTCTGAAACCAATCATTTAA	1322
Db	313	ATCTTACTCATGCACTTCCATGCAAAACAGCACATATGTTTCTGAAACCAATCATTTAA	254
Qy	1323	AGATCACAAACAGCTAGGCTTCCCGTAGCTTCCCTCTCCCTGACGATCTTTTTCG	1382
Db	253	AGATCACAAACAGCTAGGCTTCCCGTAGCTTCCCTCTCCCTGACGATCTTTTTCG	194
Qy	1383	TCACCA 1389	
Db	193	TCACCA 187	
RESULT 4			
CG224211/c		878 bp	DNA
LOCUS		OGIAG08TH ZM.0.7.1.5 KB	Zea mays genomic clone ZMMBMA0716B15,
DEFINITION			genomic survey sequence.
ACCESSION		CG224211	
VERSION		CG224211.1	GI:34124099
KEYWORDS		GSS.	
SOURCE		Zea mays	
ORGANISM		Zea mays	
		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.	
REFERENCE		1 (bases 1 to 878)	
AUTHORS		Whitefaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T., Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nuneberg,A., Robbins,D. and Lakey,N.	
TITLE		Unpublished (2002)	
JOURNAL		Other-GSSs: OGIAG08TV	
COMMENT		Contact: Cathy Whitefaw	
		TIGR	
		9712 Medical Center Drive, Rockville, MD 20850, USA	
		Tel: 301-838-5643	
		Fax: 301-838-0208	
		Email: whitefaw@tigr.org	
		Seq primer: TR	
		Class: sheared ends.	
FEATURES		location/Qualifiers	
source		1..878	
		/organism="Zea mays"	
		/mol_type="genomic DNA"	
		/strain="B73"	
		/db_xref="taxon:4577"	
		/clone_1fb="ZMMBMA0716B15"	
		/clone_1lb="ZM.0.7.1.5_KB"	
		/note="Vector: pBCK-; Site_1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"	
ORIGIN			
		4.7%: Score 66; DB 9; Length 878;	
		Best Local Similarity 100.0%; Pred. No. 5.4e-22;	
		Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1324	GATCACAAAGCTAGCGCTTCCCGCTAGCTTCCCTCTCTCTGTCGCGATCTTTTTCGT	1383
Db	878	GATCACAAAGCTAGCGCTTCCCGCTAGCTTCCCTCTCTCTGTCGCGATCTTTTTCGT	819
Qy	1384	CCACCA 1389	
Db	818	CCACCA 813	

RESULT 5
LOCUS BZ828689 539 bp DNA linear GSS 18-MAR-2003
DEFINITION PUFHM74TB ZM_0.6_1.0_KB Zea mays genomic clone ZMBMTa320N03,
genomic survey sequence.
ACCESSION BZ828689
VERSION BZ828689.1 GI:29050496
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 539)
WhiteJaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
TITLE Maize Genomics Consortium
JOURNAL Unpublished (2003)
COMMENT Other_GSSs: PUFHM74TD
Contact: Cathy WhiteJaw
TIgR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteJaw@ciqr.org
Seq primer: TR
Class: sheared ends.
FEATURES
source
1..539
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_1lb="ZMBMTa320N03"
/clone_1lb="ZM_0.6_1.0_KB"
/note="Vector: PCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
Cor selected genomic DNA library"

ORIGIN
Query Match 1.9%; Score 27; DB 8; Length 539;
Best Local Similarity 100.0%; Pred. No. 0.052;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 843 CAAAAAATTAGTTATTCTCTTT 863
|||||
DB 292 CAAAAAATTAGTTATTCTCTTT 266
|||||

RESULT 6
LOCUS BZ828694 611 bp DNA linear GSS 18-MAR-2003
DEFINITION PUFHM74TD ZM_0.6_1.0_KB Zea mays genomic clone ZMBMTa320N03,
genomic survey sequence.
ACCESSION BZ828694
VERSION BZ828694.1 GI:29050508
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 611)
WhiteJaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
TITLE Maize Genomics Consortium
JOURNAL Unpublished (2003)
COMMENT Other_GSSs: PUFHM74TB
Contact: Cathy WhiteJaw
TIgR
9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteJaw@ciqr.org
Seq primer: TR
Class: sheared ends.
FEATURES
source
1..611
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_1lb="ZMBMTa320N03"
/clone_1lb="ZM_0.6_1.0_KB"
/note="Vector: PCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
Cor selected genomic DNA library"

ORIGIN
Query Match 1.9%; Score 27; DB 8; Length 611;
Best Local Similarity 100.0%; Pred. No. 0.052;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 843 CAAAAAATTAGTTATTCTCTTT 869
|||||
DB 222 CAAAAAATTAGTTATTCTCTTT 248
|||||

RESULT 7
LOCUS CG148339 804 bp DNA linear GSS 21-AUG-2003
DEFINITION PUKB46TB ZM_0.6_1.0_KB Zea mays genomic clone ZMBMTa0780G19,
genomic survey sequence.
ACCESSION CG148339
VERSION CG148339.1 GI:34039122
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 804)
WhiteJaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
TITLE Maize Genomics Consortium
JOURNAL Unpublished (2003)
COMMENT Other_GSSs: PUKB46TD
Contact: Cathy WhiteJaw
TIgR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteJaw@ciqr.org
Seq primer: TR
Class: sheared ends.
FEATURES
source
1..804
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_1lb="ZMBMTa0780G19"
/clone_1lb="ZM_0.6_1.0_KB"
/note="Vector: PCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
Cor selected genomic DNA library"

ORIGIN
Query Match 1.9%; Score 27; DB 9; Length 804;
Best Local Similarity 100.0%; Pred. No. 0.052;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 837 TTTTTCACAAAAATTAGTTATT 863
|||||
DB 156 TTTTTCACAAAAATTAGTTATT 182
|||||

RESULT 8
CG011564/c 854 bp DNA linear GSS 19-AUG-2003
DEFINITION ZUABT23TV ZM 3.0 4.0 KB Zea mays genomic clone ZMWBPA0018C22,
genomic survey sequence.
ACCESSION CG011564
VERSION CG011564.1 GI:33883730
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 854)
AUTHORS Whiteaw,C.A., Quackenbush,J., Van Aken,S., Uteback,T.,
Reanick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
COMMENT Consortium for Maize Genomics
Unpublished (2002)
Other_GSSes: ZUABT23TV
Contact: Cathy Whiteaw
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteaw@tigr.org
Seq primer: TP
Class: sheared ends.
Location/Qualifiers
1..854
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMWBPA0018C22"
/clone_1b="ZM_3.0 4.0 KB"
/note="Vector: pBCK-; Site_1: HincII; 3-4 kb 'unfiltered'
genomic DNA library"

ORIGIN
Query Match 1.9%; Score 26; DB 9; Length 854;
Best Local Similarity 100.0%; Pred. No. 0.17; Mismatches 0; Indels 0; Gaps 0;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 268 GGGGATATATGTCCTCCACATGCTC 293
Db 221 GGGGATATATGTCCTCCACATGCTC 196

RESULT 9
BZ371637/c 687 bp DNA linear GSS 26-NOV-2002
LOCUS 1e38n03.5', genomic survey sequence.
DEFINITION BZ371637
ACCESSION BZ371637
VERSION BZ371637.1 GI:25455537
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 687)
AUTHORS Rabinowicz,P.D., O'Shaughnessy,A.L., Ballia,V., Dedhia,N.,
Katzenburger,F., King,L., Miller,B., Muller,S., Nascento,L.,
Zutavern,T., McCombie,W.R. and Martienssen,R.A.
Genomic shotgun sequences from Zea mays (methyl-filtered)
Unpublished (2002)
Contact: W. Richard McCombie
Lita Annenberg Hazen Genome Sequencing Center
Cold Spring Harbor Laboratory

PO Box 100, Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8884
Fax: 516 367 8874
Email: mccombie@cshl.org
Plate: 1e38 row: h Column: 03
Seq primer: -21M1univFwd
Class: shotgun
High quality sequence stop: 687.
Location/Qualifiers
1..687
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="1e38n03"
/lab_host="DH5a"
/clone_1b="WGS-Zmaysr (DH5a methyl filtered)"
/note="Organ: Immature ears; Site_1: Xba I; Site_2: Xba I;
The vector was digested with XbaI and one nucleotide was
added by fill in in the recessive 3' end. The genomic DNA
was nebulized, end repaired, adaptor ligated and size
fractionated using sephadex. The resulting fragments were
between 0.8 and 3 kb and were cloned into the vector (x/y
reads in M13mp19, b/g reads in pUC19). The same ligation
was transformed into DH5a."

ORIGIN
Query Match 1.8%; Score 25; DB 8; Length 687;
Best Local Similarity 100.0%; Pred. No. 0.55; Mismatches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 845 AAAAAATTAGTTATTCTCTT 869
Db 363 AAAAAATTAGTTATTCTCTT 339

RESULT 10
CC810403/c 726 bp DNA linear GSS 16-JUL-2003
LOCUS ZMWBBC0479C04r ZMWBBC Zea mays genomic clone ZMWBBC0479C04 3',
genomic survey sequence.
DEFINITION CC810403
ACCESSION CC810403
VERSION CC810403.1 GI:32817028
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 726)
AUTHORS Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C.,
Rouzaud,K., Fuks,G., Yu,Y., Wang,R. and Messing,J.
Sequencing of the maize genome at FGI (2003b)
Unpublished (2003)
Contact: Bharti,A.K.
Dr. Joachim Messing's lab
The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
University
190 Frelinghuysen Road, Piscataway, NJ 08854, USA
Tel: 732 445 3801
Fax: 732 445 5735
Email: bharti@waksman.rutgers.edu
Seq primer: Sp6
Class: BAC ends
High quality sequence start: 397.
Location/Qualifiers
1..726
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ZMWBBC0479C04"
/lab_host="E. coli DH10B"

FEATURES
source

ORIGIN	/clone_11b="ZMMB8c" /note="Vector: pTARBAC1.3; Site_1: BamHI; Site_2: BamHI"
Query Match	1.8%; Score 25; DB 9; Length 726;
Best Local Similarity	100.0%; Pred. No. 0.55;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY	265 CCAGGGATATATGCCCCACCAAT 269
Db	364 CCAGGGATATATGCCCCACCAAT 340
RESULT 11	
CG305519/c	
DEFINITION	CG305519 819 bp DNA linear GSS 25-AUG-2003 OGCGM20TV_ZM_0.7.1.5_KB_Zea mays genomic clone ZMMBMA0713C15, genomic survey sequence.
ACCESSION	CG305519
VERSION	CG305519.1 GI:34219733
KEYWORDS	GSS.
SOURCE	Zea mays
ORGANISM	Zea mays Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea. 1 (bases 1 to 819) WhiteLaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T., Reenick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N. Consortium for Maize Genomics Unpublished (2002) Other_GSSs: OG0GM20TH Contact: Cathy WhiteLaw TIGR
TITLE	9712 Medical Center Drive, Rockville, MD 20850, USA
JOURNAL	Tel: 301-838-5843
COMMENT	Fax: 301-838-0208 Email: whiteLaw@cigr.org Seq primer: TP Class: sheared ends. Location/Qualifiers 1..819 /organism="Zea mays" /mol_type="genomic DNA" /strain="B73" /db_xref="taxon:4577" /clone="ZMMBMA0713C15" /clone_11b="ZM_0.7.1.5_KB" /note="Vector: pBGSK-; Site 1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"
FEATURES	
source	
ORIGIN	
Query Match	1.8%; Score 25; DB 9; Length 819;
Best Local Similarity	100.0%; Pred. No. 0.55;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY	844 AAAAAAATTAGTTATTTCTT 868
Db	142 AAAAAAATTAGTTATTTCTT 118
RESULT 12	
CG190606/c	
LOCUS	CG190606 825 bp DNA linear GSS 21-AUG-2003 PUHHP2TB_ZM_0.6.1.0_KB_Zea mays genomic clone ZMMBTR0585M20, genomic survey sequence.
ACCESSION	CG190606
VERSION	CG190606.1 GI:34081667
KEYWORDS	GSS.
SOURCE	Zea mays
ORGANISM	Zea mays Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

REFERENCE	TITLE				JOURNAL				COMMENT			
AUTHORS	SpERMatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoidae; Andropogoneae; Zea.				1 (bases 1 to 825)				WhiteLaw,C.A., Quackenbush,J., Van Aken,S., Uteerback,T., Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and Bennezen,J.			
	Maize Genomics Consortium				Unpublished (2003)				Other_GSSs: PUHP82RD			
	Contact: Cathy Whitelaw				TIGR				9712 Medical Center Drive, Rockville, MD 20850, USA			
	Tel: 301-838-5843				Tel: 301-838-5843				Fax: 301-838-0208			
	Email: whitelaw@cigr.org				Seg primer: TR				Class: sheared ends.			
FEATURES	Location/Qualifiers				1..825				/organism="Zea mays"			
source	/mol_type="genomic DNA"				/strain="873"				/db_xref="taxon:4577"			
	/clone="ZMMBTA0585M20"				/clone_id="ZM_0.6_1.0_KB"				/note="vector: PCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high cor selected genomic DNA library"			
ORIGIN	Query Match				1.8%; Score 25; DB: 9; Length 825;				Best Local Similarity 100.0%; Pred. No. 0.55; Mismatches 0; Indels 0; Gaps 0			
Matches	25; Conservative 0; Mismatches				0; Indels				0; Gaps			
Qy	845 AAAAAATTAGTTATTTCTCTTT 869											
Db	715 AAAAAATTAGTTATTTCTCTTT 691											
RESULT 13	CG190609				833 bp				DNA			
LOCUS	PUHP82RD ZM_0.6_1.0_KB				Zea mays				genomic clone ZMMBTA0585M20,			
DEFINITION	genomic survey sequence.											
ACCESSION	CG190609											
VERSION	CG190609.1				GI:34081670							
KEYWORDS	GSS.											
SOURCE	Zea mays											
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoidae; Andropogoneae; Zea.											
REFERENCE	1 (bases 1 to 833)											
AUTHORS	WhiteLaw,C.A., Quackenbush,J., Van Aken,S., Uteerback,T., Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and Bennezen,J.											
	Maize Genomics Consortium											
	Unpublished (2003)											
TITLE	Other_GSSs: PUHP82RD											
JOURNAL	Contact: Cathy Whitelaw											
COMMENT	TIGR											
	9712 Medical Center Drive, Rockville, MD 20850, USA											
	Tel: 301-838-5843											
	Fax: 301-838-0208											
	Email: whitelaw@cigr.org											
	Seg primer: TR											
	Class: sheared ends.											
FEATURES	Location/Qualifiers											
source	1..833											
	/organism="Zea mays"											
	/mol_type="genomic DNA"											
	/strain="873"											
	/db_xref="taxon:4577"											
	/clone="ZMMBTA0585M20"											
	/clone_id="ZM_0.6_1.0_KB"											
	/note="vector: PCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high cor selected genomic DNA library"											

ORIGIN
/clone lib="ZM 0.6-1.0 KB"
/note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
COT selected genomic DNA library"

Query Match 1.8%; Score 25; DB 9; Length 833;
Best Local Similarity 100.0%; Pred. No. 0.55;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 845 AAAAAAATTAGTTATTTCTCTT 869
|||||
Db 582 AAAAAAATTAGTTATTTCTCTT 606

RESULT 14
CC358144/c 857 bp DNA linear GSS 16-MAY-2003
LOCUS
DEFINITION
CC358144
PUBPY12TD ZM 0.6-1.0 KB Zea mays genomic clone ZMMBtra507B23,
genomic survey sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
CC358144.1 GI:30827544
GSS.
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
1 (bases 1 to 857)
White, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Resnick, A., Quackenbush, J., Van Aken, S., Uteback, T.,
Bennetzen, J., Yuan, Y., San Miguel, P., Ma, J. and
Bennetzen, J.

REFERENCE
AUTHORS
White, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Resnick, A., Quackenbush, J., Van Aken, S., Uteback, T.,
Bennetzen, J., Yuan, Y., San Miguel, P., Ma, J. and
Bennetzen, J.

TITLE
JOURNAL
COMMENT
Unpublished (2003)
Other GSSs: PUBPY12TB
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.

FEATURES
source
location/Qualifiers
1..857

/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZM_0.6-1.0_KB"
/note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
COT selected genomic DNA library"

ORIGIN

Query Match 1.8%; Score 25; DB 8; Length 857;
Best Local Similarity 100.0%; Pred. No. 0.55;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 845 AAAAAAATTAGTTATTTCTCTT 869
|||||
Db 853 AAAAAAATTAGTTATTTCTCTT 829

RESULT 15
CG307825 878 bp DNA linear GSS 26-AUG-2003
LOCUS
DEFINITION
CG307825
OGMBG01TH ZM_0.7-1.5_KB Zea mays genomic clone ZMMBtra0527A01,
genomic survey sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
CG307825.1 GI:34224985
GSS.
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
1 (bases 1 to 878)
White, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Resnick, A., Quackenbush, J., Van Aken, S., Uteback, T.,
Bennetzen, J., Yuan, Y., San Miguel, P., Ma, J. and
Bennetzen, J.

REFERENCE
AUTHORS
White, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Resnick, A., Quackenbush, J., Van Aken, S., Uteback, T.,
Bennetzen, J., Yuan, Y., San Miguel, P., Ma, J. and
Bennetzen, J.

TITLE
JOURNAL
COMMENT
Unpublished (2002)
Other GSSs: OGMBG01TV
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.

FEATURES
source
location/Qualifiers
1..878

/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZMMBtra0527A01"
/note="Vector: pCR4-TOPO; Site_1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match 1.8%; Score 25; DB 9; Length 878;
Best Local Similarity 100.0%; Pred. No. 0.55;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 844 AAAAAAATTAGTTATTTCTCTT 868
|||||
Db 167 AAAAAAATTAGTTATTTCTCTT 143

Search completed: September 16, 2005, 17:48:56
Job time : 4301.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 08:25:31 ; Search time 4677.65 Seconds

(without alignments)
14440.280 Million cell updates/sec

Title: US-10-713-381-1

Perfect score: 1394
Sequence: 1 ccctggtgctctcatgaataa.....cttttcgcccaccaccatg 1394

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Genbankl:*
1: gb_ba:*
2: gb_ptg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sce:*
12: gb_sy:*
13: gb_un:*
14: gb_vt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	1394	100.0	1394	6	AX224394 Sequence
2	1394	100.0	1394	6	BD062176 Male t1ss
3	1392.4	99.9	1394	8	AF360356 Zea mays
4	1389.2	99.7	1394	6	AX224395 Sequence
5	1389.2	99.7	1394	6	BD062177 Male t1ss
6	238.8	17.1	255	6	AX224402 Sequence
7	146	10.5	158	6	AX224396 Sequence
8	77	5.5	186199	2	AC147602 Zea mays
9	73.4	5.3	25769	3	AC117267 Dictyoste
10	71.2	5.1	713	3	AF034389 Plasmodu
11	71.2	5.1	3576	3	DDU69662 Dictyoste
12	70.8	5.1	149526	2	BX957346 Danio rer
13	70.4	5.1	347582	3	PFMAL4P1 Plasmodu
14	70.2	5.0	883	6	CO422641 Sequence
15	70	5.0	241	11	BV119878 Zea mays
16	70	5.0	247	11	BV119874 Zea mays
17	70	5.0	248	11	BV119877 Zea mays
18	69.8	5.0	13400	6	CO855052 Sequence
19	69.6	5.0	139583	2	AC151050 Zea mays

C	20	69.4	5.0	110000	8	CR382132_17	Continuation (18.0
	21	69.4	5.0	181381	2	CR450818	Danio rer
	22	69.2	5.0	114817	2	AC149836	AC149836 Zea mays
	23	69.2	5.0	266544	3	AC116956	AC116956 Dictyoste
	24	69	4.9	164443	2	CR751221	CR751221 Danio rer
	25	68.8	4.9	7347	1	AF211124	AF211124 Carsonell
	26	68.6	4.9	186199	2	AC147602	AC147602 Zea mays
	27	68.4	4.9	213	11	BV119875	BV119875 PZA01377
	28	68.4	4.9	216	11	BV119871	BV119871 PZA01377
	29	68.4	4.9	224	11	BV119882	BV119882 PZA01377
	30	68.4	4.9	245	11	BV119881	BV119881 PZA01377
	31	68.4	4.9	246	11	BV119869	BV119869 PZA01377
	32	68.2	4.9	346296	8	AF090447	AF090447 Zea mays
	33	68	4.9	110000	2	PFMAL7P1_02	Continuation (3.0
	34	67.8	4.9	227495	2	AC148480	AC148480 Zea mays
	35	67.6	4.8	351	11	BV110530	BV110530 PZA01059
	36	67.6	4.8	331039	3	AC116979	AC116979 Dictyoste
	37	67.4	4.8	170	11	BV119873	BV119873 PZA01377
	38	67	4.8	2781	3	AY044085	AY044085 Dictyoste
	39	67	4.8	110000	3	AC116984_1	Continuation (2.0
	40	66.8	4.8	228	11	BV119870	BV119870 PZA01377
	41	66.8	4.8	234	11	BV119883	BV119883 PZA01377
	42	66.8	4.8	7774	12	AF338825	AF338825 Cloning v
	43	66.8	4.8	8658	12	AF338824	AF338824 Cloning v
	44	66.8	4.8	12029	3	AE001434	AE001434 Plasmodu
	45	66.8	4.8	47573	3	AF030694	AF030694 Plasmodu

ALIGNMENTS

RESULT 1	AX224394	Sequence 1 from Patent WO0160997.	1394 bp	DNA	linear	PAT 10-SBP-2001
LOCUS	AX224394					
DEFINITION	Sequence 1 from Patent WO0160997.					
ACCESSION	AX224394					
VERSION	AX224394.1	GI:15554636				
KEYWORDS						
SOURCE						
ORGANISM						
	Zea mays					
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.					
REFERENCE						
AUTHORS	Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.					
TITLE	Male tissue-preferred regulatory region and method of using same					
JOURNAL	Patent: WO 0160997-A 1 23-AUG-2001;					
PIONEER	HI-BRED INTERNATIONAL, INC. (US)					
FEATURES						
source	1.1394	/organism="Zea mays"				
		/mol_type="unasigned DNA"				
		/db_xref="taxon:4577"				
ORIGIN						
Query Match	100.0%;	Score 1394;	DB 6;	Length 1394;		
Best Local Similarity	100.0%;	Pred. No. 3.7e-270;				
Matches 1394;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;		
QY	1	CCATGCTGCTCTATGAAAAAGATGATGTCATATGTCATATCGGTTCTTCTTGAAGGCTC	60			
	1	CAATGCTGCTCTATGAAAAAGATGATGTCATATGTCATATCGGTTCTTGAAGGCTC	60			
DB	1	CAATGCTGCTCTATGAAAAAGATGATGTCATATGTCATATCGGTTCTTGAAGGCTC	60			
QY	61	CTTCTTCTGCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	120			
	61	CTTCTTCTGCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	120			
DB	61	CTTCTTCTGCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	120			
QY	121	CTTCATGTTCCACTTCTCCACCTGCGGTGACATTTCTTGAATGTCGTTGCTTCCAT	180			
	121	CTTCATGTTCCACTTCTCCACCTGCGGTGACATTTCTTGAATGTCGTTGCTTCCAT	180			
DB	121	CTTCATGTTCCACTTCTCCACCTGCGGTGACATTTCTTGAATGTCGTTGCTTCCAT	180			
QY	181	CTGACCGAGCCCATGACATCTTTCGGGACACCATCATGAGGCTTTCGATGGCCCA	240			
	181	CTGACCGAGCCCATGACATCTTTCGGGACACCATCATGAGGCTTTCGATGGCCCA	240			
DB	181	CTGACCGAGCCCATGACATCTTTCGGGACACCATCATGAGGCTTTCGATGGCCCA	240			

Db 181 CTGACGAGGCCCATGACACCTTTGCGGACACCATCAAGGCTTTGATGGCCCA 240
Qy 241 CGAGAGTATCGGGTGGTGGATCCAGGGGATATATGTCCTCCCAAAATGTCACCTATA 300
Db 241 CGAGAGTATCGGGTGGTGGATCCAGGGGATATATGTCCTCCCAAAATGTCACCTATA 300
Qy 301 TTATTTATCTTTAGATATTTATTTTGAAGAAATACAACTTATCTTTGTGTA 360
Db 301 TTATTTATCTTTAGATATTTATTTTGAAGAAATACAACTTATCTTTGTGTA 360
Qy 361 GGGCTTCAGATAGATTTTGGCTTGAAGGCCCAGAAATCGAGAGCAAGCATCTAGTG 420
Db 361 GGGCTTCAGATAGATTTTGGCTTGAAGGCCCAGAAATCGAGAGCAAGCATCTAGTG 420
Qy 421 TCCACTATGGCACTCCAGAACAGATTTAAATAAATACAAAGTAACTATCCACT 480
Db 421 TCCACTATGGCACTCCAGAACAGATTTAAATAAATAAATAAAGTAACTATCCACT 480
Qy 481 CGAAGCTATCATGTAATGTTTAAAGAACATCTATTAACACGATCCTTTAAAAA 540
Db 481 CGAAGCTATCATGTAATGTTTAAAGAACATCTATTAACACGATCCTTTAAAAA 540
Qy 541 CAAGCATATTTGAAAAGACAAATTAATGTTACAGTTTCAAAATCTAAGACGACAA 600
Db 541 CAAGCATATTTGAAAAGACAAATTAATGTTACAGTTTCAAAATCTAAGACGACAA 600
Qy 601 TTATATCGAAAGGTAAAGCTATGACGTTCAATTTTCTTTCAATCTTTATTTGTT 660
Db 601 TTATATCGAAAGGTAAAGCTATGACGTTCAATTTTCTTTCAATCTTTATTTGTT 660
Qy 661 ATTGTTTTTATATACATTTTCTTCTTACATAGATGATTTCTTCCGATTTTAAAA 720
Db 661 ATTGTTTTTATATACATTTTCTTCTTACATAGATGATTTCTTCCGATTTTAAAA 720
Qy 721 ATGACTATTAAGCATTTTATATTAAGAGACGCAATGCTAGATTTCTGTTAAAAAT 780
Db 721 ATGACTATTAAGCATTTTATATTAAGAGACGCAATGCTAGATTTCTGTTAAAAAT 780
Qy 781 TTTCTGATTTTAAAGAGCTAGTTGGCAACCTGTTCTTCAAGAAATTTGATTT 840
Db 781 TTTCTGATTTTAAAGAGCTAGTTGGCAACCTGTTCTTCAAGAAATTTGATTT 840
Qy 841 TTCAAAAAAAATTAATTTATTTTCTTTTAAATTAAGAAACATTTAGAAATAGAGT 900
Db 841 TTCAAAAAAAATTAATTTATTTTCTTTTAAATTAAGAAACATTTAGAAATAGAGT 900
Qy 901 TGGCAGACTAGCCCTAGAAATGTTTCCCAATTAATTAATCAATCTGTTATTAATTTTG 960
Db 901 TGGCAGACTAGCCCTAGAAATGTTTCCCAATTAATTAATCAATCTGTTATTAATTTTG 960
Qy 961 GCGAGCCCATTAATTAATTAACCGAACTGAATCGAGCAACCAATCTGAGCTAT 1020
Db 961 GCGAGCCCATTAATTTATTTTAAACCGAACTGAATCGAGCAACCAATCTGAGCTAT 1020
Qy 1021 TTCTAGATTAATTAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Db 1021 TTCTAGATTAATTAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Qy 1081 TGAGATGAGGAGTTTGGCAACGATAGCCAGCTATAGCTATAGTGGCTAGCTA 1140
Db 1081 TGAGATGAGGAGTTTGGCAACGATAGCCAGCTATAGCTATAGTGGCTAGCTA 1140
Qy 1141 GGTTCGAGAGCTCTCGTGCATCTACATGAGCACTACATGCTTTTCAACCGTTGCT 1200
Db 1141 GGTTCGAGAGCTCTCGTGCATCTACATGAGCACTACATGCTTTTCAACCGTTGCT 1200
Qy 1201 TTGTTTCATGTCGAAGCCTTGCTATTCTGAACCAAGAGATACCTATCCCAACAT 1260
Db 1201 TTGTTTCATGTCGAAGCCTTGCTATTCTGAACCAAGAGATACCTATCCCAACAT 1260
Qy 1261 CCATCTTACTAGCAACTTCCATGCAAGAGCACTATGTTTCTGAAACCAATCAAT 1320
Db 1261 CCATCTTACTAGCAACTTCCATGCAAGAGCACTATGTTTCTGAAACCAATCAAT 1320

Qy 1321 AAAGATCAACAAGCTAGAGCTTCCCGCTAGCTTCCCTCTCTCTGCGCATCTTTT 1380
Db 1321 AAAGATCAACAAGCTAGAGCTTCCCGCTAGCTTCCCTCTCTCTGCGCATCTTTT 1380
Qy 1381 CGTCCAGCACCATG 1394
Db 1381 CGTCCAGCACCATG 1394
RESULT 2
BD062176 1394 bp DNA linear PAT 27-AUG-2002
LOCUS BD062176
DEFINITION Male tissue-preferred regulatory region and method of using same.
ACCESSION BD062176
VERSION BD062176.1 GI:22607781
KEYWORDS JP 2001520523-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1394)
Albertsen, M.C., Fox, T.W., Garnaat, C.W., Huffman, G.A. and Kendall, T.L.
Male tissue-preferred regulatory region and method of using same
Patent: JP 2001520523-A 1 30-OCT-2001;
PIONEER HI BRED INTERNATIONAL INC
PN JP 2001520523-A/1
PD 30-OCT-2001
PF 19-JUN-1998 JP 1999504910
PR 23-JUN-1997 US 08/860499
PI MARC C ALBERTSEN, TIMOTHY W FOX, CARL W GARNAAAT, GARY A HUFFMAN,
PI TIMMY L KENDALL
PC C12N15/82, C12N15/29, C12N9/24, C12N9/22, C12N9/10, C12N9/00 PC
, C07K14/34, C12Q1/68,
PC A01H5/00
CC Strandedness: Single;
CC Topology: Linear;
FH Key location/Qualifiers.
FEATURES
source location/Qualifiers
1..1394
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 1394; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 3,7e-270;
Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CCATGAGTCTATGAAAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
Db 1 CCATGAGTCTATGAAAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
Qy 61 CTTCTTGCCTTATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 120
Db 61 CTTCTTGCCTTATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 120
Qy 121 CTCATGTTCCACTTCTCCCACTCGCGTTGACATTTCTTGGATGAGTGGTGGTCCAT 180
Db 121 CTCATGTTCCACTTCTCCCACTCGCGTTGACATTTCTTGGATGAGTGGTGGTCCAT 180
Qy 181 CTGACGAGGCCCATCAAGCACTTTCGAGACACCATCAAGGCTTTGAGATGGCCCA 240
Db 181 CTGACGAGGCCCATCAAGCACTTTCGAGACACCATCAAGGCTTTGAGATGGCCCA 240
Qy 241 CGAGAGTATCGGGTGGTGGATCCAGGGGATATATGTCCTCCCAAAATGTCACCTATA 300
Db 241 CGAGAGTATCGGGTGGTGGATCCAGGGGATATATGTCCTCCCAAAATGTCACCTATA 300
Qy 301 TTATTTATCTTTAGATATTTATTTTGAAGAAATACAACTTATCTTTGTGTA 360
Db 301 TTATTTATCTTTAGATATTTATTTTGAAGAAATACAACTTATCTTTGTGTA 360

QY 361 GGGCTCAGCATAGATTTCGCTTAGGCCCCGAAATGCCAGACCAAGCCATGCTAGTG 420
 DB 361 GGGCTCAGCATAGATTTCGCTTAGGCCCCGAAATGCCAGACCAAGCCATGCTAGTG 420
 QY 421 TCCACTATGGCACTACCCAGAAACAAAGATTAAATAAATAACCAAGTAACTAATCCACT 480
 DB 421 TCCACTATGGCACTACCCAGAAACAAAGATTAAATAAATAACCAAGTAACTAATCCACT 480
 QY 481 CGAAAGCTATCATGTATGTTTAAAGAAACATCTATTAAACACGATCCCTTTAAAAAA 540
 DB 481 CGAAAGCTATCATGTATGTTTAAAGAAACATCTATTAAACACGATCCCTTTAAAAAA 540
 QY 541 CAAGCATATTTCCGAAAGACAAATTTGTTACAGTTTACAAACATTTAAGAGACAAA 600
 DB 541 CAAGCATATTTCCGAAAGACAAATTTGTTACAGTTTACAAACATTTAAGAGACAAA 600
 QY 601 TTAATATGAAAGGTAAGCTATGACGTTGAGATTTCCTTTTCATTTCTGTTATTTTGT 660
 DB 601 TTAATATGAAAGGTAAGCTATGACGTTGAGATTTCCTTTTCATTTCTGTTATTTTGT 660
 QY 661 ATTGTTTTATATACATTTTCTCTCTTACAAATAGATGATTTTCTCCGATTTTAA 720
 DB 661 ATTGTTTTATATACATTTTCTCTCTTACAAATAGATGATTTTCTCCGATTTTAA 720
 QY 721 ATGACTATAAGCATTTTATATATAGACACGATGCTGATTTCTCGTTCAAAATC 780
 DB 721 ATGACTATAAGCATTTTATATATAGACACGATGCTGATTTCTCGTTCAAAATC 780
 QY 781 TTTCTGATTTTTTAAGAGTAGTTTGGCAACCTGTTCTTCAAGAAATTTGATTTT 840
 DB 781 TTTCTGATTTTTTAAGAGTAGTTTGGCAACCTGTTCTTCAAGAAATTTGATTTT 840
 QY 841 TTCAAAAAAATAGTTTATTTTCTCTTATATAAATAGAAAACATTTAGAAAATAGAGT 900
 DB 841 TTCAAAAAAATAGTTTATTTTCTCTTATATAAATAGAAAACATTTAGAAAATAGAGT 900
 QY 901 TGCAGACTAGCCCTAGAAATGTTTCCCAATAAATTCATCACTGTTATTAATTTG 960
 DB 901 TGCAGACTAGCCCTAGAAATGTTTCCCAATAAATTCATCACTGTTATTAATTTG 960
 QY 961 GCCAGCCCCATTAATTTTAAACCGAAATCGAAATGAGCGAAACCAATTCGAGCTAT 1020
 DB 961 GCCAGCCCCATTAATTTTAAACCGAAATCGAAATGAGCGAAACCAATTCGAGCTAT 1020
 QY 1021 TTCTAGATTAGTAAAGAGAGAGAGAGAGAGAAATCAGTTTAACTGATTTGCC 1080
 DB 1021 TTCTAGATTAGTAAAGAGAGAGAGAGAGAGAAATCAGTTTAACTGATTTGCC 1080
 QY 1081 TGAAGATGTCGGGTTTGGCAACGATAGCCAGCTATATAGCTCATAGTGCTTACGTCA 1140
 DB 1081 TGAAGATGTCGGGTTTGGCAACGATAGCCAGCTATATAGCTCATAGTGCTTACGTCA 1140
 QY 1141 GGTTCGGACCTCTCGTGCATCTCAATGCGCATACATCAATGCTGTTCAACGGTCTC 1200
 DB 1141 GGTTCGGACCTCTCGTGCATCTCAATGCGCATACATCAATGCTGTTCAACGGTCTC 1200
 QY 1201 TTGTTCCATGTCGCAAGCCTTGCCTATTCGAAACCAAGAGATACCTATCCCAACAT 1260
 DB 1201 TTGTTCCATGTCGCAAGCCTTGCCTATTCGAAACCAAGAGATACCTATCCCAACAT 1260
 QY 1261 CCATCTTACTGACGAACTTCATGCAACGACACATATGTTTCCGAAACCAATCATTT 1320
 DB 1261 CCATCTTACTGACGAACTTCATGCAACGACACATATGTTTCCGAAACCAATCATTT 1320
 QY 1321 AAAGATCAGAAAGCTAGGTTCTCCCGCTAGGTTCCCTCTCTGCTGCGAATCTTTT 1380
 DB 1321 AAAGATCAGAAAGCTAGGTTCTCCCGCTAGGTTCCCTCTCTGCTGCGAATCTTTT 1380
 QY 1381 CGTCAACCAACATG 1394
 DB 1381 CGTCAACCAACATG 1394
 QY 1394 CGTCAACCAACATG 1394
 DB 1394 CGTCAACCAACATG 1394

RESULT 3
 AF360356 3343 bp DNA linear PLN 12-MAY-2001
 LOCUS
 DEFINITION Zea mays male fertility protein (M45) gene, complete cds.
 ACCESSION AF360356
 VERSION AF360356.1 GI:14028756
 KEYWORDS
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 3343)
 Fox,T.W., Trimmell,M.R. and Albertsen,M.C.
 Cloning of M45, a gene required for male fertility from Zea mays
 Unpublished
 2 (bases 1 to 3343)
 Fox,T.W., Trimmell,M.R. and Albertsen,M.C.
 Direct Submission
 Submitted (13-MAR-2001) Trait and Technology Development, Pioneer
 Hi-Bred Inc., Inc., 7300 N.W. 62nd Ave., P.O. Box 1004, Johnston,
 IA 50131-1004, USA
 Location/Qualifiers
 source
 1..3343
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /culturvar="B73"
 /db_xref="taxon:4577"
 /chromosome="9L"
 <1392..>2942
 /gene="M45"
 join(<1392..1768,1898..2182,2280..2447,2534..>2942)
 /gene="M45"
 /product="male fertility protein"
 join(1392..1768,1898..2182,2280..2447,2534..2942)
 /gene="M45"
 /codon_start=1
 /product="male fertility protein"
 /protein_id="AAK52489.1"
 /db_xref="GI:14028757"
 /translation="MEKRNLOWRGRDGIVOYPHLFPALALALLVADPGSLPAEY
 DYPVKNHEALPYGVNGSPRDNAERLGRGLFEYGEVDEPSTEDDLOGRPYAGLA
 DGRVVRWMEGEAGWETPAVMNPDSVEVCANVNSTRRQHEKEFCGPIGLRFG
 TGEIVYDAYVGLAVVGGGVAASVAREADDPPIRFANDLVRHNGSVFETDSNRK
 SRKDHNLILIEEGTGLRLRYPETSGVIVLKGIVPAGVQISEDHQFLPSETNRC
 RIMRYWLEGPRAGEVFPANLPSPDNVNSNGQFWVADICRTPDQDFVAPKPLR
 TLVFKPLSLIKYLTKWAKARMHTVIALLDGEKRVVLEDRHEHMKLVSEVRVGRK
 LMIGVANHNTIIPLED"
 ORIGIN
 Query Match 99.9%; Score 1392.4; DB 8; Length 3343;
 Best Local Similarity 99.9%; Pred. No. 7.5e-270;
 Matches 1393; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CCATGCTCTCTATGAAAAAGATGACAAATGCTCTATATCCGTTTCTTAGGGTCC 60
 DB 1 CCATGCTCTCTATGAAAAAGATGACAAATGCTCTATATCCGTTTCTTAGGGTCC 60
 QY 61 CTTCTTCGCTTATTAAGCACTGAATCGGGGTTTCAAAAACTTCCAGGGGTGATGAT 120
 DB 61 CTTCTTCGCTTATTAAGCACTGAATCGGGGTTTCAAAAACTTCCAGGGGTGATGAT 120
 QY 121 CTCATGTTTCACTTCTCCACCTCGCTTGACATTTCTTGAGATGCGGTGATCCCAT 180
 DB 121 CTCATGTTTCACTTCTCCACCTCGCTTGACATTTCTTGAGATGCGGTGATCCCAT 180
 QY 181 CTGACCGAGCCCATCAGACACTTTGCGGACACCCATCAGAGGCTTTGCGATGGCCCA 240
 DB 181 CTGACCGAGCCCATCAGACACTTTGCGGACACCCATCAGAGGCTTTGCGATGGCCCA 240
 QY 241 CGAGACGTATGGGTCGTGATGATCAGGGGATATATGTCGCCCAATTCGTCACCTATA 300
 DB 241 CGAGACGTATGGGTCGTGATGATCAGGGGATATATGTCGCCCAATTCGTCACCTATA 300

Db	241	CGAAGCATATCGGGTGGTGGATCCAGGGGATATATGCCCCCAAGATGTCACCTATA	300
Qy	301	TTTATATTCTTTAGATTTTATTAATTTTGGAAAAATAACAATTATATCTTTGTGA	360
Db	301	TTTATATTCTTTAGATTTTATTAATTTTGGAAAAATAACAATTATATCTTTGTGA	360
Qy	361	GGGCTTCAGCATGATTTTGGCTTAGGGCCAGAAATGGAGGACCGACGATGCTAGTG	420
Db	361	GGGCTTCAGCATGATTTTGGCTTAGGGCCAGAAATGGAGGACCGACGATGCTAGTG	420
Qy	421	TCCACTATTGGCACTACCCAGAAACAAGATTTAAAAAATAACAAAGTAACTATCCACT	480
Db	421	TCCACTATTGGCACTACCCAGAAACAAGATTTAAAAAATAACAAAGTAACTATCCACT	480
Qy	481	CGAAGCTATCATGATATGTTTAAAGAAATCTATTTAAAAACACGATCCTCTTAAAAA	540
Db	481	CGAAGCTATCATGATATGTTTAAAGAAATCTATTTAAAAACACGATCCTCTTAAAAA	540
Qy	541	CAAGCATATTTCCAAAAGACAATAATGTTACAGTTACAAACATCTAAGACGACAAA	600
Db	541	CAAGCATATTTCCAAAAGACAATAATGTTACAGTTACAAACATCTAAGACGACAAA	600
Qy	601	TTATATCGAAAAGGTAAAGCTATGACGTTCAAGATTTTCTTTTCATTCTGTATATTTGTT	660
Db	601	TTATATCGAAAAGGTAAAGCTATGACGTTCAAGATTTTCTTTTCATTCTGTATATTTGTT	660
Qy	661	ATTGTTTTTATACATTTTCTTCTCTTACATAGATGATTTTCTTCCGATTTTATATA	720
Db	661	ATTGTTTTTATACATTTTCTTCTCTTACATAGATGATTTTCTTCCGATTTTATATA	720
Qy	721	ATGACTATAAAGTATTTTATATATAAGAGCAGCAGTCGTAGATTCGTCAAAAATC	780
Db	721	ATGACTATAAAGTATTTTATATATAAGAGCAGCAGTCGTAGATTCGTCAAAAATC	780
Qy	781	TTTCTGATTTTTTAAAGAGCTAGTTGGCAACCCCTGTTCTTTCAAAAGAAATTTGATTT	840
Db	781	TTTCTGATTTTTTAAAGAGCTAGTTGGCAACCCCTGTTCTTTCAAAAGAAATTTGATTT	840
Qy	841	TTTCAAAAAAATAGTTTATTTCTCTTTTAAATAAGAAACAATTAGAAATAGAGT	900
Db	841	TTTCAAAAAAATAGTTTATTTCTCTTTTAAATAAGAAACAATTAGAAATAGAGT	900
Qy	901	TGCGAGCTAGCCCTAGAAATGTTTCCCAATAAATTACAATCACTGTATTAATTATTTG	960
Db	901	TGCGAGCTAGCCCTAGAAATGTTTCCCAATAAATTACAATCACTGTATTAATTATTTG	960
Qy	961	GCCAGCCCCATTAATTTTAAACCGAACTGAATCGAGCGAAACCAAACTGAGCTAT	1020
Db	961	GCCAGCCCCATTAATTTTAAACCGAACTGAATCGAGCGAAACCAAACTGAGCTAT	1020
Qy	1021	TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAATCAATTTTAAGTCAATTGCC	1080
Db	1021	TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAATCAATTTTAAGTCAATTGCC	1080
Qy	1081	TGAGATGTGGGTTTGGCAACGATAGCAGACCGTAACTCATAGCTCATAGGTCCTAGCTCA	1140
Db	1081	TGAGATGTGGGTTTGGCAACGATAGCAGACCGTAACTCATAGGTCCTAGCTCA	1140
Qy	1141	GGTTGGGAGCTCTCGTGTATCTCACATGGCAACTACATGCTTGTTCACACGTTGCTC	1200
Db	1141	GGTTGGGAGCTCTCGTGTATCTCACATGGCAACTACATGCTTGTTCACACGTTGCTC	1200
Qy	1201	TTGTTCCATGCTGCAAGCCCTGCTATTCTGAACCAAGAGATACCTAACCAAACTAT	1260
Db	1201	TTGTTCCATGCTGCAAGCCCTGCTATTCTGAACCAAGAGATACCTAACCAAACTAT	1260
Qy	1261	CCATCTTACTCATGCACTTCATGCAACACGCAATATGTTTCCGAAACCAATCCATT	1320
Db	1261	CCATCTTACTCATGCACTTCATGCAACACGCAATATGTTTCCGAAACCAATCCATT	1320
Qy	1321	AAAGATCAACAAGCTAGCGTTCTCCGCTAGCTTCCCTCTCTCTCTGCGCATCTTTT	1380
Db	1321	AAAGATCAACAAGCTAGCGTTCTCCGCTAGCTTCCCTCTCTCTCTGCGCATCTTTT	1380

Qy	1381	CGTCCACCAACCATG	1394
Db	1381	CGTCCACCAACCATG	1394
RESULT 4			
AX224395			
LOCUS	AX224395	1394 bp	DNA linear PAT 10-SEP-2001
DEFINITION	Sequence 2 from Patent WO0160997.		
ACCESSION	AX224395		
VERSION	AX224395.1	GI:15554637	
KEYWORDS			
SOURCE			
ORGANISM	Zea mays		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.		
REFERENCE			
AUTHORS	1		
TITLE	Albertsen, M.C., Fox, T.W., Garnat, C.W., Huffman, G. and Kendall, T.L.		
JOURNAL	Male ribosome-preferred regulatory region and method of using same		
FEATURES			
source	1. 1394		
	/organism="Zea mays"		
	/mol_type="unassigned DNA"		
	/db_xref="caxon:4577"		
ORIGIN			
Query Match	99.7%; Score 1389.2; DB 6; Length 1394;		
Best Local Similarity	99.8%; Pred. No. 3.4e-269;		
Matches 1391; Conservative	0; Mismatches 3; Indels 0; Gaps 0;		
Qy	1	CCATGCTGCTCTATGAAAAAGATGATCAATGTGTCTATATCCGTTTCTTAGGGTCC	60
Db	1	CCATGCTGCTCTATGAAAAAGATGATCAATGTGTCTATATCCGTTTCTTAGGGTCC	60
Qy	61	CTTCTTGGCCTTATTAAGTGAATGGGGTTTCAAAAACTTCCAGGGTGCATGAT	120
Db	61	CTTCTTGGCCTTATTAAGTGAATGGGGTTTCAAAAACTTCCAGGGTGCATGAT	120
Qy	121	CTCCATGTTTCACTTCTCCCACTCGCGTGGACATTTCTTGGATGTCGTTGCCAT	180
Db	121	CTCCATGTTTCACTTCTCCCACTCGCGTGGACATTTCTTGGATGTCGTTGCCAT	180
Qy	181	CTGACCGAGGCCCATCAACACCTTTGCGGACACCCATCAAGGCGCTTTGGATGGCCA	240
Db	181	CTGACCGAGGCCCATCAACACCTTTGCGGACACCCATCAAGGCGCTTTGGATGGCCA	240
Qy	241	CGAGACGTATCGGGTCGTGGTGAATCCAGGGATATATGTCCTCCCAATCGTCACTATA	300
Db	241	CGAGACGTATCGGGTCGTGGTGAATCCAGGGATATATGTCCTCCCAATCGTCACTATA	300
Qy	301	TTATTAATCTTTAGATATTTAATTTTGGAAAAATAACAATTATTAATCTTTGTGA	360
Db	301	TTATTAATCTTTAGATATTTAATTTTGGAAAAATAACAATTATTAATCTTTGTGA	360
Qy	361	GGGCTTCAGCATGATTTTGGCTTAGGGCCAGAAATGGAGGACCGACGATGCTAGTG	420
Db	361	GGGCTTCAGCATGATTTTGGCTTAGGGCCAGAAATGGAGGACCGACGATGCTAGTG	420
Qy	421	TCCACTATTGGCACTACCCAGAAACAAGTAAATTAACCAAGTAACTATCCACT	480
Db	421	TCCACTATTGGCACTACCCAGAAACAAGTAAATTAACCAAGTAACTATCCACT	480
Qy	481	CGAAGCTATCATGATATGTTTAAAGAAATCTATTTAAAAACACGATCCTCTTAAAAA	540
Db	481	CGAAGCTATCATGATATGTTTAAAGAAATCTATTTAAAAACACGATCCTCTTAAAAA	540
Qy	541	CAAGCATATTTCCAAAAGACAATAATGTTTACAGTTTACAAACATCTAAGACGACAAA	600
Db	541	CAAGCATATTTCCAAAAGACAATAATGTTTACAGTTTACAAACATCTAAGACGACAAA	600

QY 601 TTATATCGAAGGTAAGCTATGAGCTCAGATTTTCTTTTCAATCTTGTATTTGTT 660
DB 601 TTATATCGAAGGTAAGCTATGAGCTCAGATTTTCTTTTCAATCTTGTATTTGTT 660
QY 661 ATGTGTTTATATACATTTTCTTCTCTTACAAATGAGTGAATTTCTCCGATTTTAA 720
DB 661 ATGTGTTTATATACATTTTCTTCTCTTACAAATGAGTGAATTTCTCCGATTTTAA 720
QY 721 ATGACTATAAAGCTATTTTATATAGAGACGAGATGCTAGATTTCTCGTTCAAAATC 780
DB 721 ATGACTATAAAGCTATTTTATATAGAGACGAGATGCTAGATTTCTCGTTCAAAATC 780
QY 781 TTTCTGATTTTATTAAGAGCTAGTTGGCAACCTGTTCTTCAAGAAATTTGATTT 840
DB 781 TTTCTGATTTTATTAAGAGCTAGTTGGCAACCTGTTCTTCAAGAAATTTGATTT 840
QY 841 TTCAAAAAATAGTTATTTTCTCTTATTAATAAATGAAAAACATTAAGAAATAGACT 900
DB 841 TTCAAAAAATAGTTATTTTCTCTTATTAATAAATGAAAAACATTAAGAAATAGACT 900
QY 901 TGGCAGACTAGCCCTAGAAATGTTTCCCAATAATTCAATCACTGTATTAATTTTG 960
DB 901 TGGCAGACTAGCCCTAGAAATGTTTCCCAATAATTCAATCACTGTATTAATTTTG 960
QY 961 GCCAGCCCCATTAATTTTAAACCGAAATCGAAATCGAGCGAAACCAATTCGAGCTAT 1020
DB 961 GCCAGCCCCATTAATTTTAAACCGAAATCGAAATCGAGCGAAACCAATTCGAGCTAT 1020
QY 1021 TTCTCTAGATTAGTAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
DB 1021 TTCTCTAGATTAGTAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
QY 1081 TGAAGATGTCGGGTTTGGCAACGATAGCCAGTATCATATGATGATGATGATGATGATG 1140
DB 1081 TGAAGATGTCGGGTTTGGCAACGATAGCCAGTATCATATGATGATGATGATGATGATG 1140
QY 1141 GGTTCGGAGACTCTCGTGTATCTCACATGAGCAATGCTGTTTCAACCGTTCTGTC 1200
DB 1141 GGTTCGGAGACTCTCGTGTATCTCACATGAGCAATGCTGTTTCAACCGTTCTGTC 1200
QY 1201 TTGTTTCATGTCGTCAGAGCTTGGCTATTTGAAACCAAGAGATACCTACGCCAACAAT 1260
DB 1201 TTGTTTCATGTCGTCAGAGCTTGGCTATTTGAAACCAAGAGATACCTACGCCAACAAT 1260
QY 1261 CCATCTTACTCATGCAACTTCATGCAACACGACATATGTTTCCGAAACAGATCATTT 1320
DB 1261 CCATCTTACTCATGCAACTTCATGCAACACGACATATGTTTCCGAAACAGATCATTT 1320
QY 1321 AAAGATCAACAAGCTAGGCTTCTCCGCTAGCTTCCCTCTCTCTGCGCATTTT 1380
DB 1321 AAAGATCAACAAGCTAGGCTTCTCCGCTAGCTTCCCTCTCTCTGCGCATTTT 1380
QY 1381 CGTCCACCAACGATG 1394
DB 1381 CGTCCACCAACGATG 1394

RESULT 5
LOCUS BD062177 1394 bp DNA linear PAT 27-AUG-2002
DEFINITION Male tissue-preferred regulatory region and method of using same.
ACCESSION BD062177.1 GI:22607782
VERSION JP 2001520523-A/2.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1394)
AUTHORS Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G.A. and
Kendall,T.L.
TITLE Male tissue-preferred regulatory region and method of using same

JOURNAL Patent: JP 2001520523-A 2 30-OCT-2001;
PIONEER HI BRED INTERNATIONAL INC
COMMENT PN JP 2001520523-A/2
PD 30-OCT-2001
PF 19-JUN-1998 JP 199504910
PR 23-JUN-1997 US 08/880499
PI MARC C ALBERTSEN, TIMOTHY W FOX, CARL W GARNAT, GARY A HUFFMAN,
TIMMY L KENDALL
PC C12N15/82, C12N15/29, C12N9/24, C12N9/22, C12N9/10, C12N9/00 PC
PC C07K14/34, C1201/68,
PC A01H5/00
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers.
FEATURES
source 1..1394
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 99.7%; Score 1389.2; DB 6; Length 1394;
Best Local Similarity 99.8%; Pred. No. 3,4e-269;
Matches 1391; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CCATGCTGCTCTATGAAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 60
DB 1 CCATGCTGCTCTATGAAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 60
QY 61 CTTCCTCTGCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 120
DB 61 CTTCCTCTGCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 120
QY 121 CTCATGTCATCACTTCTCCACCTCGCTGTCATATTTCTGATGATGATGATGATGATGATG 180
DB 121 CTCATGTCATCACTTCTCCACCTCGCTGTCATATTTCTGATGATGATGATGATGATGATG 180
QY 181 CTGACCGAGGCCCATCAACACCTTTCCGGAACACCATTAAGAGGCTTTCCGATGGCCCA 240
DB 181 CTGACCGAGGCCCATCAACACCTTTCCGGAACACCATTAAGAGGCTTTCCGATGGCCCA 240
QY 241 CGAGAGCTATGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
DB 241 CGAGAGCTATGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
QY 301 TTATTAATCTTATGATATTAATTTTGGAAAAATTAACAACCTTAATTAATTTTGTGTA 360
DB 301 TTATTAATCTTATGATATTAATTTTGGAAAAATTAACAACCTTAATTAATTTTGTGTA 360
QY 361 GGGCCTCAGCATGATGATTTTCCGCTAGGGCCGAAATGAGAGGACAGGCAATGCTAGTG 420
DB 361 GGGCCTCAGCATGATGATTTTCCGCTAGGGCCGAAATGAGAGGACAGGCAATGCTAGTG 420
QY 421 TCCACATATGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
DB 421 TCCACATATGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
QY 481 CGAAAGCTATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
DB 481 CGAAAGCTATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
QY 541 CAAGCATATTTGAAAGAGACAAATTAATGATGATGATGATGATGATGATGATGATGATG 600
DB 541 CAAGCATATTTGAAAGAGACAAATTAATGATGATGATGATGATGATGATGATGATGATG 600
QY 601 TTATATCGAAGGTAAGCTATGAGCTCAGATTTTCTTTTCAATCTTGTATTTGTT 660
DB 601 TTATATCGAAGGTAAGCTATGAGCTCAGATTTTCTTTTCAATCTTGTATTTGTT 660
QY 661 ATGTGTTTATATACATTTTCTTCTCTTACAAATGAGTGAATTTCTCCGATTTTAA 720
DB 661 ATGTGTTTATATACATTTTCTTCTCTTACAAATGAGTGAATTTCTCCGATTTTAA 720

```

Qy 721 ATGACTATAAAGTCATTTTATATATAGAGCAGCATGTCAGATTCTGTCACAAAATC 780
    |||||
Db 721 ATGACTATAAAGTCATTTTATATATAGAGCAGCATGTCAGATTCTGTCACAAAATC 780
Qy 781 TTTCTGATTTTTTTTAAAGAGCTAGTTTGGCAACCCGTGTTCTTCAAGAAATTTGAATTT 840
    |||||
Db 781 TTTCTGATTTTTTTTAAAGAGCTAGTTTGGCAACCCGTGTTCTTCAAGAAATTTGAATTT 840
Qy 841 TTTCAAAAAAATTAAGTTTATTTTCTCTTTATATATATATATATATATATATATATAT 900
    |||||
Db 841 TTTCAAAAAAATTAAGTTTATTTTCTCTTTATATATATATATATATATATATATATAT 900
Qy 901 TGGCAGACTAGCCCTAGAAATGTTTTCCCAATTAATTCATCACTGTTATTAATTTT 960
    |||||
Db 901 TGGCAGACTAGCCCTAGAAATGTTTTCCCAATTAATTCATCACTGTTATTAATTTT 960
Qy 961 GCCAGCCCCCAATTAATTTTAAACCGAATCGAAATCGAAGAACCAATCTGAGCTAT 1020
    |||||
Db 961 GCCAGCCCCCAATTAATTTTAAACCGAATCGAAATCGAAGAACCAATCTGAGCTAT 1020
Qy 1021 TTCTCTAGATTAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
    |||||
Db 1021 TTCTCTAGATTAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Qy 1081 TGAGATGTGCGGTTTGGCAAGATAGCCACCGTATATCATAGTGTGCTAGCTCA 1140
    |||||
Db 1081 TGAGATGTGCGGTTTGGCAAGATAGCCACCGTATATCATAGTGTGCTAGCTCA 1140
Qy 1141 GGTTCGGCAGCTCTCGTGCATCTCACATGGCATACTACATGCTTTGTTCAACCGTTGCT 1200
    |||||
Db 1141 GGTTCGGCAGCTCTCGTGCATCTCACATGGCATACTACATGCTTTGTTCAACCGTTGCT 1200
Qy 1201 TTGTTCCATGTCGCAAGCCCTGCTTATTTGAACCAAGAGATACCTATCCCAACAAAT 1260
    |||||
Db 1201 TTGTTCCATGTCGCAAGCCCTGCTTATTTGAACCAAGAGATACCTATCCCAACAAAT 1260
Qy 1261 CCATCTTACTCATGCACTTCCATGCAACACGACATATGTTTCCGTAACCAATCATTT 1320
    |||||
Db 1261 CCATCTTACTCATGCACTTCCATGCAACACGACATATGTTTCCGTAACCAATCATTT 1320
Qy 1321 AAAGATCAACAACAGCTTAGCTTCTCCGCTAGCTTCCCTCTCTCTGCGCATCTTTT 1380
    |||||
Db 1321 AAAGATCAACAACAGCTTAGCTTCTCCGCTAGCTTCCCTCTCTCTGCGCATCTTTT 1380
Qy 1381 CGTCCACCCACCATG 1394
    |||||
Db 1381 CGTCCACCCACCATG 1394

```

```

RESULT 6
AX224402 255 bp DNA linear PAT 10-SEP-2001
LOCUS AX224402
DEFINITION Sequence 9 from Patent WO0160997.
ACCESSION AX224402
VERSION AX224402.1 GI:15554644
KEYWORDS
SOURCE
ORGANISM Zea mays
            Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1
AUTHORS Albertsen,M.C., Fox,T.W., Garnat,C.W., Huffman,G. and Kendall,T.L.
TITLE Male tissue-preferred regulatory region and method of using same
JOURNAL Patent: WO 0160997-A 9 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
    source
        1..255
        /organism="Zea mays"
        /mol_type="unassigned DNA"
        /db_xref="taxon:4577"
ORIGIN

```

```

Query Match 17.1%; Score 238.8; DB 6; Length 255;
Best Local Similarity 97.2%; Pred No. 8.3e-38;
Matches 243; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy 1145 CGGAGCTTCGTCATCTCATGCAATGCAATGCTTGTTCACCGTTGCTTGT 1204
    |||||
Db 5 CCGGAGATCCCGTCATCTCATGCAATGCAATGCTTGTTCACCGTTGCTTGT 64
Qy 1205 TCCATGTCACAGCTTGGCTTATTTGAAACCAAGAGATACCTACTTCCCAACATTCAT 1264
    |||||
Db 65 TCCATGTCACAGCTTGGCTTATTTGAAACCAAGAGATACCTACTTCCCAACATTCAT 124
Qy 1265 CTTACTCATGCAACTTCCATGCAACACGACATATGTTTCTGTAACCAATCATTAAG 1324
    |||||
Db 125 CTTACTCATGCAACTTCCATGCAACACGACATATGTTTCTGTAACAGATCTATTAAAG 184
Qy 1325 ATCACAACAGCTAGCGTTCTCCGCTAGCTTCCCTCTCTCTGCGCATCTTTTTCGTC 1384
    |||||
Db 185 ATCACAACAGCTAGCGTTCTCCGCTAGCTTCCCTCTCTCTGCGCATCTTTTTCGTC 244
Qy 1385 CACCAACCATG 1394
    |||||
Db 245 CACCAACCATG 254

```

```

RESULT 7
AX224396 158 bp DNA linear PAT 10-SEP-2001
LOCUS AX224396
DEFINITION Sequence 3 from Patent WO0160997.
ACCESSION AX224396
VERSION AX224396.1 GI:15554638
KEYWORDS
SOURCE
ORGANISM Zea mays
            Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1
AUTHORS Albertsen,M.C., Fox,T.W., Garnat,C.W., Huffman,G. and Kendall,T.L.
TITLE Male tissue-preferred regulatory region and method of using same
JOURNAL Patent: WO 0160997-A 3 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
    source
        1..158
        /organism="Zea mays"
        /mol_type="unassigned DNA"
        /db_xref="taxon:4577"
ORIGIN
Query Match 10.5%; Score 146; DB 6; Length 158;
Best Local Similarity 99.4%; Pred No. 3.9e-19;
Matches 157; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
Qy 1155 CGTGCATCTCATGCAATGCAATGCTTGTTCACCGTTGCTC-TTGTTCATGCTC 1213
    |||||
Db 1 CGTGCATCTCATGCAATGCAATGCTTGTTCACCGTTGCTC-TTGTTCATGCTC 60
Qy 1214 CAAGCTTGGCTTATTTGAAACCAAGAGATACCTACTTCCCAACCAATCATTTTCTCAT 1273
    |||||
Db 61 CAAGCTTGGCTTATTTGAAACCAAGAGATACCTACTTCCCAACCAATCATTTTCTCAT 120
Qy 1274 GCAACTTCATGCAACACGACATATGTTTCTGTAAC 1311
    |||||
Db 121 GCAACTTCATGCAACACGACATATGTTTCTGTAAC 158

```

ORIGIN

KEYWORDS
SOURCE
ORGANISM

HTG; HTGS_PHASE2; HTGS_FULLTOP; HTGS_ACTIVEFIN.

Zea mays

Eukaryotes; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

1 (bases 1 to 186199)

Bjren, B., Nusbaum, C., Lander, E., Butler, E., Wing, R., Bharti, A.K.

and Messing, J.

Zea mays, clone ZMNBC0334A01

Unpublished

TITLE
JOURNAL
REFERENCE
AUTHORS

2 (bases 1 to 186199)

Bjren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,

Anderson, M., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T.,

Bogulavsky, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y.,

Collamore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K.,

Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,

Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S.,

Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hages, B.,

Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,

Kamat, A., Karacas, A., Kells, C., Landers, T., Levine, R.,

Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., Maclean, C.,

Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,

Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J.,

Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,

O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,

Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Riese, C., Rogov, P.,

Roman, J., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C.,

Talamas, J., Teefaye, S., Theodore, J., Topham, K., Travers, M.,

Vasilev, H., Venkataraman, V., Viel, R., Vo, A., Wilson, B., Wu, X.,

Wymann, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (19-DEC-2003) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 186199)

Bjren, B., Nusbaum, C., Lander, E., Butler, E., Wing, R., Bharti, A.K.,

Messing, J., Abouelleil, A., Allen, N., Bastien, V., Bloom, T.,

Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Bogulavsky, L.,

Boukhalter, B., Camarata, J., Chang, J., Choepel, Y., Collamore, A.,

Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J.S., Dodge, S.,

Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P.,

Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L.,

Grand-Pierre, N., Hafez, N., Hagopian, D., Hages, B., Hall, J.,

Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,

Karacas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,

Liu, G., Liu, X., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P.,

Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J.,

Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,

Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,

O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,

Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Riese, C., Rogov, P.,

Roman, J., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C.,

Spencer, B., Stange-Thomann, N., Stojanovic, N., Subb, M.,

Talamas, J., Teefaye, S., Theodore, J., Topham, K., Travers, M.,

Vasilev, H., Venkataraman, V., Viel, R., Vo, A., Wilson, B., Wu, X.,

Wymann, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (17-AUG-2004) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Aug 17, 2004 this sequence version replaced gi:4658659.

All repeats were identified using RepeatMasker:

Shit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/MIT Center for Genome Research

Center code: WITR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@road.mit.edu

Bharti, A.K. and Messing, J.: The Plant Genome Initiative at

Rutgers, Maksum Institute, Rutgers, The State University of New

Jersey, 190 Frelinghuysen Road, Piscataway, NJ 08854, USA

(http://pgir.rutgers.edu)

Butler, E and Wing, R: Arizona Genomics Institute, Biological

Sciences West, 448A, P.O. Box 210088, University of Arizona,

Tucson, AZ 85721, USA (http://www.genome.arizona.edu)

Project Information

Center project name: L30003

Center clone name: 334_A_1

Consensus Information

This consensus is derived from a shotgun assembly that has been

manually curated. It is the best representation of the BAC that we

can generate without further laboratory work. The draft assembly

has been edited, and if possible, ends identified by vector as well

as by BAC end sequences, and contigs ordered and oriented. Bases

that are not Ns are either above Q20 or manually edited. This

assembly was performed with Arachne (Genome Res. 2002 12: 177-189;

Genome Res. 2003 13: 91-96). All trace files for this project are

available at the NCBI trace repository

(http://www.ncbi.nlm.nih.gov/Traces/trace.fcgi?l). An exact list of

reads used in this assembly are available at

http://www.broad.mit.edu/annotation/plants/maize/randomclones.html.

NOTE: This is a 'working draft' sequence. It currently

consists of 6 contigs. Gaps between the contigs

are represented as runs of N. The order of the pieces

is believed to be correct as given, however the sizes

of the gaps between them are based on estimates that have

provided by the submitter.

This sequence will be replaced

by the finished sequence as soon as it is available and

the accession number will be preserved.

1 100617: contig of 100617 bp in length

100618 100717: gap of unknown length

100718 104730: contig of 4013 bp in length

104731 104830: gap of unknown length

104831 115104: contig of 10274 bp in length

115105 115204: gap of unknown length

115205 156386: contig of 41192 bp in length

156387 156496: gap of unknown length

156497 179936: contig of 23440 bp in length

179937 180036: gap of unknown length

180037 186199: contig of 6163 bp in length.

Location/Qualifiers

1..186199

/organism="Zea mays"

/mol_type="genomic DNA"

/cultivar="B73"

/db_xref="taxon:4577"

/clone="ZMNBC0334A01"

/clone_lib="CHORI-201 Maize B73 BAC MDO1 Library"

ORIGIN

Query Match 5.5%; Score 77; DB 2; Length 186199;

Best Local Similarity 71.4%; Pred. No. 2.3e-05;

Matches 115; Conservative 0; Mismatches 45; Indels 1; Gaps 1;

Db 162058

840

162117

900

162177

AC117267

AC117267

AC117267

AC117267

AC117267

AC117267

AC117267

AC117267

ACCESSION AC117267 AC115597
VERSION AC117267.2 GI:42733680
KEYWORDS
SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
REFERENCE Eukaryota; Mycetozoa; Dictyostelid; Dictyostelium.
AUTHORS 1 (bases 1 to 25769)
Gloeckner, G., Eichinger, L., Szefranski, K., Pachbat, J., Dear, P.,
Lehmann, R., Baumgart, C., Parra, G., April, J. F., Gligo, R., Kumpf, K.,
Tunggal, B., Cox, E., Quail, M. A., Platzer, M., Rosenthal, A. and
Noegel, A. A.
TITLE Sequence and analysis of chromosome 2 of Dictyostelium discoideum
JOURNAL Nature 418 (6833), 79-85 (2002)
MEDLINE 22092622
PUBMED 12097910
REMARK The Dictyostelium Genome Sequencing Consortium
AUTHORS 2 (bases 1 to 25769)
Baumgart, C.
TITLE Direct Submission
JOURNAL Submitted (09-APR-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
REFERENCE 3 (bases 1 to 25769)
Baumgart, C.
TITLE Direct Submission
JOURNAL Submitted (21-FEB-2004) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
COMMENT On or before Feb 21, 2004 this sequence version replaced
gi:19570016, gi:20087114.
CDS predictions from GenBank may contain errors. Further information
is available from IMB Jena, Department of Genome Analysis
(http://genome.imb-jena.de/dictyostelium/)
and the University Cologne, Institute for Biochemistry 1
(http://www.uni-koeln.de/dictyostelium/project.html)
Punding
Agency : Deutsche Forschungsgemeinschaft (DFG).
Location/Qualifiers
1. 25769
/organism="Dictyostelium discoideum"
/mol_type="genomic DNA"
/strain="AX4"
/db_xref="taxon:44689"
/chromosome="2"
/map="5836255-5862024"
/compement (join(832..1070,1237..1349,1533..1798,
1896..2053,2231..2360))
/note="Geneid exon scores (in order of location ranges) :
23.68, 1.54, 15.75, 9.21, -4.75 - GSCJ_ID dd_03387"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAS38625.1"
/db_xref="GI:42733681"
/translation="MRLQILFLSTILLIPOLKSPSSNDISSTILSTPNPNKDVNNO
ILSKSEFRSLSESSNKRYIFVFDGPNHEIENLRITKEPKNKYGVDSFSFCL
NSKFNLEVKRSTISRNKVPENIMWIESTHCYNNSVYLNKNNNEVEY
FSQVAKSFLQNGYNIKISSKNTFKIFEENNDKLFSSLPFYLNKIMENDEFEY
SSPLGWTDFKIKQFKLNDIETILSDIEVNLCSNNNNNNNNVDETLTLGSSGSG
SGSGSSSGSGAGQISGTGYVNHQ"
join(12957..3173,3272..5583)
/note="Geneid exon scores (in order of location ranges) :
24.88, 254.96 - GSCJ_ID dd_00753"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAS38626.1"
/db_xref="GI:42733682"
/translation="MEPPTKYVYITSNLDPKPKFPOSEMSLTTTDTIOKSFVNSG
SSSSSKGFGEAVLDLDTAGQEYSAMRQVYRTDCFMIVSIDSRSFEEVSLK
QHIERVADRDVPIIIVGNKVDESRRQVRIADQLASLRVPIETSAKRSNIER
AFETLVHTRPNTVYKVVGGGVSALIIQFIQNHFEVDEPTIEDSKRVQVYIS
GPIPIGSLNKSSSSSSSSSSSGGKLENKIFSGDKQSPQQAASSTIDRTGCI
STRLFANLYSVSMNLSEKVPILITGDVYCGCCVYLISRFSLAKTGPDSFWKGF
CKKSNITLLEGEIIPKDSVEYVLSPTSSITSDSKKBEETIICIVDSGAGIT
EVSJLSEWNAKGVGASGPSYISRLCVSSYIPMIDRLSIQYPKRVLVLTFS
DEVMIYTSNVAQPIVIAQDKLEDFDLIEGRSMTYDCLPTASGSSDFLKAKIKSL

EPVOSTALGPALLVSAALASOKMLSEVICTDGVNVLGAIEDLPLGPAQFEYKVT
KLAONKTTITIGISGSHIDLGIVGKSEONNGSITTIHPIELAREIRKLPONMIA
TDVMSICLHPLEINKYDSKQGLSRVYKQFPNNNSITDITLLYSRRRPTFOVQY
FOIQIKYTDGVRCLRVASAOQATPPNTSISANISILMAFTQOAKTAQOEY
MESRLHLKATKTLIRSLCNTDQWEEFNVFEVLEMEKAPLITCIKNQOVRKATD
DSIRVFKWKNKHFVSGRKKDISRRRGAEBINKOYNIKFT"
complement(5711..6796)
/note="Geneid exon scores (in order of location ranges) :
68.61 - GSCJ_ID dd_00752"
/codon_start=1
/product="similar to Plasmodium falciparum (isolate 3D7).
Hypothetical protein"
/protein_id="AAS38627.1"
/db_xref="GI:42733683"
/translation="MSSRVKSLIERVNGINLTIEKSLSPRGILNTNINNNNNNN
GTTTTSAGANNNSSTPSTVPVDFRVHSSNNQIOIYVDSFPNTNNNNNN
NN
KNNNNNNYFDFSTNEMITTLNSCTLEQGTIDNCYSTNTPSPSSSTSLAN
KSNHMMMPSLKINOIINDNNNNININFEVQISNSSSLSSGCTYQSCPIDEHIOE
QLRIQEKQKOIDDODDDDDDDDDSDSPPPPTTTTKTFCLSPRIKKEENID
QNSMPPPKLDRRRSRDLIPTANI"
join(8924..9076,9236..10573)
/note="Geneid exon scores (in order of location ranges) :
10.37, 128.60 - GSCJ_ID dd_00729"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAS38628.1"
/db_xref="GI:42733684"
/translation="MSSSTITPIATTTTNNNSSTNGHNSKTNFTLEFKNIKNINPS
ENNAIKNSVGIENGPNVNNNNDRKLSNDGVNFGKNLTPPSKNSMSTKHS
SSSSDLKANEDYESIKELKMSDITGCSPLSPRESSDITIDESTATTK
LIDAVAAASSSSSVKTYTIDVLSFKSNKTKPIODIYENHOKIYVLSPTTPYN
NNNNNNNNSSOHTANQOKYPIESQIISPKFIAYAOSSSTKTNINNNNNNTT
TKKANSITTPGSNNRNHYNNTSKSKKKNPIPSLNTATPNQNTITPYSKSTT
TPKSNKINDVNAAPAAVANSITTEBPVALITDATTSTNSPKTPEKEDTPKLA
AAQROIDIGKTAGYKNYIALVPKSRKRPDTPKNQOVCSKRSMDQIKKMRQLH
ENYDQIDIPNEVEASNESTESTLEKEMNKLEISENENHN"
complement(10972..11979)
/note="Geneid exon scores (in order of location ranges) :
23.29 - GSCJ_ID dd_00728"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAS38629.1"
/db_xref="GI:42733685"
/translation="MEKIKFYEFLIDFONOKFKIOBFTSKLIGKESFTTPEPIVY
EYLNWQTSIESILKNGTINKSIFBIFSVCGYIGELREYRIVELKPTTPIIL
DYIRKINSLWKNKEKLNLSLITNNKNEKIGFBWEIINEFONFTKESFOR
ITNANSIMIKSIDSRNISNSPINSRCSFNEFEKRSDDAGDFTLITSMGIMONYI
ENSIISEKFPOTINHAKSPFLINDLVSPNEIENBDLANTIKLIAIOLNSIOESID
KTEELIIDHYLKLSSIFILKLYQNDSTYQLKQVONSKILISGIYFAHKKSKRY
N"
join(12884..13020,13172..15320)
/note="Geneid exon scores (in order of location ranges) :
2.42, 136.56 - GSCJ_ID dd_00727"
/codon_start=1
/product="similar to Dictyostelium discoideum (slime
mold). Hypothetical 97.7 kDa protein"
/protein_id="AAS38630.1"
/db_xref="GI:42733686"
/translation="MDNLIYDFLIKEKNQGSFLDCILLSINGLYISYDGLDYCNN
FLINENKVAKSAISYGLANVOYQLDLEFRLKLNKNTDCKRKINTHNSLKKCC
GNKPDYGLIIIEENNNIINDQYFNNKNSNNNIYNNNDNNNSIIEEDLNKXI
TLKNERINGDEITTTTTTTTSSSTTMMTTTNGVNRNVLTPQVARNYVIK
TOLIFHLALYNHAKKYFLTPICLVYKFKCYLOSMTLHNDDDADDESYSDI
DDDSYSDGNCSSSSGSDIGSSNSINNGISSSSSILNSNDLSMTIRVNSL
BGLKSIIEFEKENVYIQRLPSSISIRFSGFQRLAKGVISNVSITSGDSFNO
SLDGKMSLKQHLHLOFGHFOQITMGQPLSILSLIILDPKCYVIERGSLPDSVKT
LDYKNSLQNSOISIFNPIPKSITLVDSBENONIKANDISNVLISIFGSEFNS
DIGISLNSIIEIKGRAPRDIACSSITSIDFGKKRNPPLSMOTLRSIDSGS
KNGOITPQGITHTLKLNSRGYHNOIIPADLTPYLESLNLGYNREITYKNDYD
CYGINKGFGSSNSNFCVGGTNGHLEMLKNTSLKTLTILNFKRKEIVDDLPSI
ESLNGYHNPDIQNNVLPKLLKQLFLNSEFNONISADGCPFLGLQTIYIRNSMNF
INSLNPLFTIKYINIIDLHL"

Oy		504	AAGAAACATCTATTAAAAACGAGCCCTTTAAAAAACAAGACTTTTCGAAAGACAGA	563
Dd		621	AATAATTATCCTATGTATTAATTAAGCATATGTGGAAATTTATTTTTTATGGAAAATATAAAT	562
Oy		564	ATTATGTTACAGTTTACAAACATCTAAGACGCAAAATTATATCGAAAGGTAAAGCTATGA	623
Dd		561	ATAATGTAAATATTTTTTTTTTTTATTAATATA-AATATAATATATATATATATTTTTTT	503
Oy		624	CGTTCAGATTTTTCTTTTTCATTTCTGTATTTTGTTATGTGTTTTTATATACATTTTCTT	683
Dd		502	TTTAAAA	443
Oy		684	CTCTTACATAGAGTATTTTCTCCGATTTTATATAAATGACTATAAGTCATTTTATTA	743
Dd		442	TTTTCAGTATATAAAGTATTTTAACGGAAATATGTGAATGCAAAAAGAAATTTGTATG	383
Oy		744	TAAAGCACGCATGTCGTATGATTTCTGTTCAAAAATCTTTCTGATTTTTTAAAGCTAG	803
Dd		382	TAAATTTTAAATATAAATAATAAAAAGAAATTAATTTTTTTATATAATATATAAAG	323
Oy		804	TTTGGCAACCCTGTTCTTTCAAGAAATTTGATTTTTTTCAAAAAATTAAGTTATATT	863
Dd		322	TTTATTTATTTTATATATATATATTTTTTTTTTTAAATTTCAATATATTTTAAATTT	263
Oy		864	CTCTTTATATAAATATAGAAAAACATTAGAAAAATAGATGTCGACAGTACCCCTAGAAATG	923
Dd		262	ATATTTTTTTTTTTTTTTTTTGTCTATATATATATATATATAATACGTAAATATATATATAT	203
Oy		924	TTCCCAATTAATTTACATCATCTGTATATATTTTGGCCAGCCCCCATTAATTTATTTAA	983
Dd		202	TTTTTTTTTTTTTTTTTTTAAATTTTGAAATAATATTTATTTTCTTCAATATATATAATTA	143
Oy		984	CCGAACATGAATTCAGCAAAACCAATCTGACTATTTCTTACATTAGTAAAAAGGA	1043
Dd		142	TAAATTTTATATTTTAAAGCTATCATTTAATGTTTTTTTTPAAAAATAAAAAAAAAATAGT	83
Oy		1044	GAGAGAGAGAGAGANA	1059
Dd		82	TATCTATATATPAAAA	67

RESULT 11
DDU86962/c

LOCUS DDU86962 3576 bp DNA linear INV 23-JUL-2001

DEFINITION Dictyostelium discoideum miga (miga) gene, complete cds.

ACCESSION U86962

VERSION U86962.1 GI:1841871

KEYWORDS

SOURCE ORGANISM Dictyostelium discoideum
Dictyostelium discoideum
Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
1 (bases 1 to 3576)
Escalante, R., Messels, D., Soll, D.R. and Loomis, W.F.
Chemotaxis to cAMP and slug migration in Dictyostelium both depend on miga, a BTB protein
Mol. Biol. Cell 8 (9), 1763-1775 (1997)

JOURNAL MEDLINE
PUBMED 9307972
2 (bases 1 to 3576)
Escalante, R., Messels, D., Soll, D. and Loomis, W.F.
Direct Subdivision
Submitted (27-JAN-1997) Department of Biology, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-032, USA

Db 612 TTGTTATTTTGTATTTTAAATTAATTTATTAACATTAATAAATCTTTAAACACAC 553

Qy 712 TTTTATTAATAATGACTATAAGTCATTTTATATTAAGACAGCATGTCGATTCGT 771

Db 552 ACACACACACACCAATATGTTTTTTTTTGTGTTTTTGTGTTTTTTTATCAACCT 493

Qy 772 TCAAAATCTTTCGATTTTAAAGCAGTATGTCGACACCTGTTCTTTCAAGAT 831

Db 492 TTTCAATCGCTTTGATTTAGTTTGTGTTTTTTTTTTTGAATGTTTTTTGAAAAAT 433

Qy 832 TTTGATTTTTCAAAAAATAGTTATTTCTCTTAT 871

Db 432 ATTTTATTTTATTTATTTTATTTTATTTTATTTTATTTT 393

RESULT 12
BX957346 149526 bp DNA linear HTG 08-OCT-2004
LOCUS Danio rerio clone CH211-117K16, WORKING DRAFT SEQUENCE.
DEFINITION BX957346.13 GI:54019944
VERSION HTG; HTGS_PHASE2; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS Danio rerio (zebrafish)
SOURCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
ORGANISM 1 (bases 1 to 149526)
REFERENCE Beasley, H.
AUTHORS Direct Submission
JOURNAL Submitted (07-OCT-2004) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Oct 9, 2004 this sequence version replaced gi:53850295.

COMMENT ----- Genome Center.
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: zc117K16
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 149526 bases at least Q40
Consensus quality: 149526 bases at least Q30
Consensus quality: 149526 bases at least Q20
Insert size: 149526; sum-of-config
Quality coverage: 12.41x in Q20 bases; sum-of-configs Quality coverage: 12.54x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES
source
1. 149526
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="CH211-117K16"
/clone_1ib="CHORI-211"
1. 149526
/note="assembly_fragment:03115
clone_end:SP6
vector_side:right
clone_end:T7
vector_side:left"

misc_feature
1. 149526
/note="assembly_fragment:03115
clone_end:SP6
vector_side:right
clone_end:T7
vector_side:left"

ORIGIN
Query Match 5.1%; Score 70.8; DB 2; Length 149526;
Best Local Similarity 46.8%; Pred. No. 0.00041;

Matches 260; Conservative 0; Mismatches 292; Indels 4; Gaps 1;

Qy 488 TATCATGTATGTTTAAAGAAACATCTATTAACACGATCCTTAAAAACAGCAT 547

Db 109063 TTTTAATATATGAACAAAATCTATTTTACTTTTGCCCTTGTGATGATAGTCA 109004

Qy 548 ATTTCGAAGAGACAAATATATGTTACA---GTTTCAAAACATCTAAGCGACAAATTA 603

Db 109003 GTTTTAAAGATTTAAAAATTTTAAACATGTTTATGATCATATATAGCTAAAGAGA 108944

Qy 604 TATCGAAGGTATAGTATGATGACGTTGATTTTCTTTTCATCTGTTATTTGTTAT 663

Db 108943 CCAAAAGTACTAATTAGCATTTTAAATGTTTATTTTATTTTATTTTATTT 108884

Qy 664 GTTTTATATACATTTTCTTCTTTCAATAGATGATTTCTTCGATTTTATTAATAATG 723

Db 108883 AT 108824

Qy 724 ACTATTAAGTCAATTTTATATTAAGACAGCATGTCGTGATTTCTGTTCAAAATCTTT 783

Db 108823 TATATATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 108764

Qy 784 CTGATTTTATTAAGAGCTAGTTTGCAACCTGTTCTTTCAAGAAATTTGATTTTTC 843

Db 108763 ATATATCATATTTAAACCTTTATATGTTAATGTTAATTAATTAATTAATTAATGTTCA 108704

Qy 844 AAAAAAATTAAGTTTATTTTCTTTTATTAATAATGAACCACTAGAAAAATAGAGTTGC 903

Db 108703 AGAAAAACAAAGTTAATTTTATTAATCTTATTTTATTAATTAATTAATTAATTAATTTT 108644

Qy 904 CAGACAGCCCTGAGATGTTTCCCAATTAATTAATTAATTAATTAATTAATTTTGGC 963

Db 108643 TATTTATGCTTCACTTACTGATCTTTATTTTAAATTAATTAATTAATTAATTTTGC 108584

Qy 964 AGCCCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1023

Db 108583 ACATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 108524

Qy 1024 TCTAGATTAGTAATAA 1039

Db 108523 GTTTTATATGTTAATAA 108508

RESULT 13
PFMAL4P1 347582 bp DNA linear INV 29-JAN-2003
LOCUS Plasmodium falciparum MAL4P1.
DEFINITION AL034557.8 GI:23498126
ACCESSION AL034557.8 GI:23498126
VERSION
KEYWORDS Plasmodium falciparum 3D7
SOURCE Plasmodium falciparum 3D7
ORGANISM Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE 1
AUTHORS Hall, N., Pain, A., Berriman, M., Churcher, C., Harris, B., Harris, D., Mungall, K., Bowman, S., Atkin, R., Baker, S., Barron, A., Brooks, K., Buckee, C.O., Burrows, C., Cherevach, I., Chillingworth, C., Chillingworth, T., Christodoulou, Z., Clark, L., Clark, R., Corcoran, C., Cronin, A., Davies, R., Davis, P., Dear, P., Dearden, F., Doggett, J., Felwell, T., Goble, A., Goodhead, I., Gilliam, R., Hamlin, N., Hane, Z., Harper, D., Hauser, H., Hornby, T., Holroyd, S., Horrocks, P., Humphrey, S., Jagels, K., James, K.D., Johnson, D., Keriohou, A., Knights, A., Kontoriotis, B., Kyes, S., Larke, N., Lawson, D., Leonard, N., Line, A., Maddison, M., McLean, J., Mooney, P., Moulé, S., Murphy, L., Oliver, K., Ormond, D., Price, C., Quail, M.A., Rabinowitsch, E., Rajandream, M.A., Rutter, S., Rutherford, K.M., Sanders, M., Simmonds, M., Seeger, K., Sharp, S., Smith, R., Squares, R., Squares, S., Stevens, K., Taylor, K., Tivey, A., Unwin, L., Whitehead, S., Woodward, J., Sulston, J.E., Craig, A., Newbold, C. and Barrell, B.G.
TITLE Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13
JOURNAL Nature 419 (6906), 527-531 (2002)
MEDLINE 22255708

Db 273 NTYT 270

RESULT 15
BV119878/c 241 bp DNA linear STS 19-MAR-2004
P2A01377 CML247 Zea mays CML247 Zea mays STS genomic, sequence
tagged site.

ACCESSION BV119878
VERSION BV119878.1 GI:45589251
KEYWORDS STS.
SOURCE Zea mays
ORGANISM Zea mays

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
1 (bases 1 to 241)
McMullen, M.D., Vroch Bi, I., Schroeder, S.S. and Gaut, B.S.
MPZ-UCI Joint SNP Discovery
Unpublished (2003)

COMMENT

Contact: Brandon S. Gaut
Dept. Ecology and Evolutionary Biology
U.C. Irvine
321 Steinhilber Hall, Irvine, CA 92697-2525, USA
Tel: (949) 824-2564
Fax: (949) 824-2181
Email: bgaut@uci.edu
Primer A: TTCTACGTCGAATGATGATGAGA
Primer B: AGGAACCAAGCTTATGAGAGA
STS size: 241
Protocol:
PCR amplification of genomic DNA
Template: 50 ng
Primer: each 0.5 uM
dNTPs: each 200 uM
Taq Polymerase: RedTaq (Sigma)
Total Vol: 10 uL
Amplicon sequencing
ABI protocol - using d-Rhodamine terminator cycle
Sequencing ready reaction with ampliTaq DNA polymerase FS
Sequencing ran on ABI 3700 sequencer.

Buffer:
Genomic DNA amplification
RedTaq (Sigma)
Sequencing buffer
d-Rhodamine Kit (ABI)

PHRED/PRAP Quality Scores 56 56 56 51 51 40 40 40 40 35 35 42 48
56 56 56 56 42 42 56 33 33 33 47 44 37 37 37 37 48 48
42 42 42 42 30 30 30 37 37 44 44 84 84 90 89 81 71 68 51 40 45
45 51 39 39 56 44 48 86 86 96 81 71 68 63 62 62 58 58 69 69
67 67 77 72 68 71 68 72 74 80 74 76 73 66 77 79 68 62 57 56 53 60
58 71 71 68 68 77 75 79 71 79 68 81 75 90 79 35 37 33 39 42 42 35
37 30 38 20 48 40 44 40 33 42 37 74 68 60 50 54 54 51 51 55
49 50 57 57 72 82 90 88 88 79 74 74 67 64 62 53 50 52 59 55
65 71 71 84 84 84 88 85 90 84 88 88 90 80 79 83 83 86 86 81
76 68 49 48 39 39 42 49 33 42 29 24 24 14 14 14 32 44 40 48 48 90
84 84 84 87 79 72 72 63 57 53 51 41 57 54 56 56 49 51 30 33 33 28
28 28 33 40 42 48 37 37.

FEATURES
source

Location/Qualifiers
1..241
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="CML247"
/db_xref="taxon:4577"
/clone_lib="Zea mays CML247"
/dev_stage="seedling"
/note="Organ: leaf; genomic DNA from inbred line"
<1..>241

STS
ORIGIN

Query Match 5.0%; Score 70; DB 11; Length 241;
Best Local Similarity 61.5%; Pred. No. 0.00074;
Matches 112; Conservative 0; Mismatches 70; Indels 0; Gaps 0;
Qy 799 GCTAGTTGGCAACCTGTTCTTTCAAGATTGATTTTCAAAAAAATTAGTTT 858
Db 195 GTTAGTTGGAAACACATTTTCCCAAGGAAATTTCCCAAGAAATTAAGTTC 136
Qy 855 ATTTTCTCTTTAATAAATGAAAAACATTAGAAAAATAGCTGGCAGACTAGCCCTAGA 918
Db 135 ATTTTCCTTGGAAAAATGAAATCCCAAGGAAAAATGAGTTCCCAACTAACCTTAAT 76
Qy 919 ATGTTTCCCAATTAATTACATCTGTGTATTAATTATTTGGCCAGCCCAATTAATAT 978
Db 75 ACTGTTAATTATTTAGTTCCCAATATCCATCACCAAGATGATATCTTAACCGCATATCTGT 16
Qy 979 TT 980
Db 15 TT 14

Search completed: September 16, 2005, 03:01:11
Job time : 4684.65 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 08:10:17 ; Search time 674.892 Seconds
(without alignments)
12227.321 Million cell updates/sec

Title: US-10-713-381-1

Perfect score: 1394
Sequence: 1 cccatgctgcctcatgaaaa.....cttttcgcccaccaccatg 1394

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: geneseqn16dec04:*
2: geneseqn19908:*
3: geneseqn20008:*
4: geneseqn2001a8:*
5: geneseqn2001b8:*
6: geneseqn2002a8:*
7: geneseqn2002b8:*
8: geneseqn2003a8:*
9: geneseqn2003b8:*
10: geneseqn2003c8:*
11: geneseqn2003d8:*
12: geneseqn2004a8:*
13: geneseqn2004b8:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1394	100.0	1394	2	AAx07408	Aax07408 Zea mays
2	1394	100.0	1394	2	AAH76332	Aah76332 Z. mays M
3	1389.2	99.7	1394	2	AAx07409	Aax07409 Zea mays
4	1389.2	99.7	1394	5	AAH76333	Aah76333 Z. mays M
5	238.8	17.1	255	5	AAH76340	Aah76340 Z. mays M
6	146	10.5	158	5	AAH76334	Aah76334 Z. mays M
7	70.2	5.0	883	4	AA15210	Aa15210 Human bre
8	70.2	5.0	960	11	ACN85231	Acn85231 Breast ca
9	69.8	5.0	13400	13	ADRo4286	AdRo4286 Corn FT h
10	67.2	4.8	2657	2	AAZ10551	Aaz10551 DNA sequ
11	66.4	4.8	6027	2	AAx58751	Aax58751 Maize dul
12	66.4	4.8	6027	8	ABX09335	Abx09335 DNA encod
13	66.4	4.8	6027	12	ADK12106	Adk12106 cDNA enc
14	66.2	4.7	439	8	ABX35844	Abx35844 Bovine ES
15	65.6	4.7	346	4	AA187279	Aa187279 Human pol
16	65.6	4.7	8056	8	ABZ10246	Abz10246 Haematopo
17	64.6	4.6	9859	13	ADs89714	Ads89714 Oligonuc1
18	64.4	4.6	9859	13	ADs89440	Ads89440 Oligonuc1
19	64	4.6	8056	8	ABZ10100	Abz10100 Haematopo
20	63.8	4.6	431	8	ABX44556	Abx44556 Bovine ES

C	21	63.6	4.6	335913	5	AA161371	Aa161371 Soybean 2
C	22	63.6	4.6	335913	5	AA161372	Aa161372 Soybean 2
C	23	63.4	4.5	12968	4	AA545494	Aa545494 Chemicall
C	24	63.4	4.5	12968	4	AA546779	Aa546779 Tumour su
C	25	63.4	4.5	12968	6	AB134085	Ab134085 Human imm
C	26	63.4	4.5	12968	6	ABK28424	Abk28424 DNA trans
C	27	63.4	4.5	12968	6	ABN80285	Abn80285 Human che
C	28	63.4	4.5	12968	6	ABN80285	Abn80285 Human che
C	29	62.4	4.4	10762	6	AB134113	Ab134113 Human imm
C	30	61.4	4.4	9652	6	AB192332	Ab192332 Chemicall
C	31	60.6	4.3	565	13	ACN56344	Acn56344 Cotton an
C	32	60.6	4.3	5822	6	AB133097	Ab133097 Human imm
C	33	60.4	4.3	500	5	AD173089	Ad173089 Human ova
C	34	60.4	4.3	500	5	AD173089	Ad173089 Human ova
C	35	60.2	4.3	5818	6	AA546636	Aa546636 Tumour su
C	36	60.2	4.3	8896	6	ABN80326	Abn80326 Human che
C	37	59.4	4.3	461	5	AD143834	Ad143834 Human ova
C	38	59.4	4.3	14919	4	AA546506	Aa546506 Tumour su
C	39	59.2	4.2	19459	6	ABK31213	Abk31213 Signal tr
C	40	59.2	4.2	19459	6	ABL70528	Ab170528 Chemicall
C	41	59	4.2	517	13	ACN56273	Acn56273 Cotton an
C	42	59	4.2	5303	6	AB132871	Ab132871 Human imm
C	43	58.8	4.2	393	8	ABX39417	Abx39417 Bovine ES
C	44	58.6	4.2	392	5	AD143653	Ad143653 Human ova
C	45	58.6	4.2	529	13	ACN56029	Acn56029 Cotton an

ALIGNMENTS

RESULT 1
AAx07408 standard; DNA; 1394 BP.
XX
XX
AC AAX07408;
XX
DT 08-JUN-1999 (first entry)
XX
DE Zea mays Me45 male tissue-preferred regulatory region.
XX
DE Me45; male; tissue-preferred; regulatory region; plant cells;
KM plant tissue; differentiated; maize; hybrid seed; fertility; ss.
XX
OS Zea mays.
XX
PN WO9859061-A1.
XX
PD 30-DEC-1998.
XX
PF 19-JUN-1998; 98WO-US012895.
XX
PR 23-JUN-1997; 97US-00880499.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman GA, Kendall TL;
XX
DR WPI; 1999-105628/09.
XX
PT New nucleic acid encoding a Me45 male tissue-preferred regulatory region
PT - useful in mediating plant fertility, especially hybrid seed production.
XX
XX Claim 2; Page 22-23; 39pp; English.
XX
XX The sequence is that encoding an Me45 male tissue-preferred regulatory
XX region. It may be used in the construction of a vector for a method of
XX producing exogenous genes in a male tissue-preferred manner, which is
XX useful in restoring or conferring fertility, such as in hybrid seed
XX production. In conferring fertility, a monocot/dicot plant is transformed
XX with the exogenous nucleotide sequence (a male sterility gene, preferably
XX Me45), which encodes a product selected from auxins, rolls and diptheria
XX toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
XX and infertile plants

XX Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 1394; DB 2; Length 1394;
Best Local Similarity 100.0%; Pred. No. 5.1e-291;
Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCATGGGTCTCTATGAAAAAGATGATACATGTCTATATCCGTTTTCTTAGGGTCC 60
DB 1 CCATGGGTCTCTATGAAAAAGATGATACATGTCTATATCCGTTTTCTTAGGGTCC 60
QY 61 CTTCTTCTGCTTATTACTGACTGAATCGGGGTTTACAAAAAATCTTCACGGGTGCATGAT 120
DB 61 CTTCTTCTGCTTATTACTGACTGAATCGGGGTTTACAAAAAATCTTCACGGGTGCATGAT 120
QY 121 CTGCATTTCCACTTCCCACTCGGGTTGCACTTTCTTGATGTGCGGTGTTCCCAT 180
DB 121 CTGCATTTCCACTTCCCACTCGGGTTGCACTTTCTTGATGTGCGGTGTTCCCAT 180
QY 181 CTGACCGAGGCGCATCGACACCTTCGGGACACCCATCAAGGGCCTTTGGATGGCCCA 240
DB 181 CTGACCGAGGCGCATCGACACCTTCGGGACACCCATCAAGGGCCTTTGGATGGCCCA 240
QY 241 CGAGACGTATCGGGTGTGTGATCGAGGGATATATGTCCCCCAATCGTCACTATA 300
DB 241 CGAGACGTATCGGGTGTGTGATCGAGGGATATATGTCCCCCAATCGTCACTATA 300
QY 301 TTATATCTTCTTGATATTTTAAATTTTGGAAAAATTAACAACCTATATCTTTGTGTA 360
DB 301 TTATATCTTCTTGATATTTTAAATTTTGGAAAAATTAACAACCTATATCTTTGTGTA 360
QY 361 GGGCCCTCAGATAGATTTTGGCTTAGGGCCAGAAATGCGAGACACGCATGCTAGTG 420
DB 361 GGGCCCTCAGATAGATTTTGGCTTAGGGCCAGAAATGCGAGACACGCATGCTAGTG 420
QY 421 TCCACTATTTGGCACTACCCAGAACAGATTTAAAAAATAACCAAGTATCAATCACT 480
DB 421 TCCACTATTTGGCACTACCCAGAACAGATTTAAAAAATAACCAAGTATCAATCACT 480
QY 481 CGAAAGCTATCAGTATATGTTTAAAGAAATCATATTAACCAACGATCCTCTTAAAAA 540
DB 481 CGAAAGCTATCAGTATATGTTTAAAGAAATCATATTAACCAACGATCCTCTTAAAAA 540
QY 541 CAAGCATATTTCCGAAAGACAAATTAATGTTTACAGTTTACAAACATCTAGACGCAAA 600
DB 541 CAAGCATATTTCCGAAAGACAAATTAATGTTTACAGTTTACAAACATCTAGACGCAAA 600
QY 601 TTATATGAAAGGTAGCTATGACGTTCAAGATTTTCTTTTCAATCTTGTAATTTGTT 660
DB 601 TTATATGAAAGGTAGCTATGACGTTCAAGATTTTCTTTTCAATCTTGTAATTTGTT 660
QY 661 ATTGTTTTTATATACATTTCTTCTCTTACATATGAGTATTTCTTCGATTTTAAAA 720
DB 661 ATTGTTTTTATATACATTTCTTCTCTTACATATGAGTATTTCTTCGATTTTAAAA 720
QY 721 ATGACTATAAGTCATTTTATATATAGAGACCGCATGCTAGATTTCTCGTTCAAAAATC 780
DB 721 ATGACTATAAGTCATTTTATATATAGAGACCGCATGCTAGATTTCTCGTTCAAAAATC 780
QY 781 TTTCTGATTTTTTTAGAGCTAGTTTGCAACCCGTGTTCTTCAAGAAATTTGATTTT 840
DB 781 TTTCTGATTTTTTTAGAGCTAGTTTGCAACCCGTGTTCTTCAAGAAATTTGATTTT 840
QY 841 TTGCAAAAAAATAGTTTATTTTCTCTTTTAAAAATGAAAAACATTAAGAAAAATGAGT 900
DB 841 TTGCAAAAAAATAGTTTATTTTCTCTTTTAAAAATGAAAAACATTAAGAAAAATGAGT 900
QY 901 TGGCAGACTAGCCCTAGAAATGTTTTCCCAATAATTAACATCACTGTATTAATTTTG 960
DB 901 TGGCAGACTAGCCCTAGAAATGTTTTCCCAATAATTAACATCACTGTATTAATTTTG 960
QY 961 GCCAGCCCTAATAATTTTAAACCGAACTGAATCGACGAAACCAATCTGACCTAT 1020
DB 961 GCCAGCCCTAATAATTTTAAACCGAACTGAATCGACGAAACCAATCTGACCTAT 1020

DB 961 GCCAGCCCTAATAATTTTAAACCGAACTGAATCGACGAAACCAATCTGACCTAT 1020
QY 1021 TTCTCTAGATTAGTATAAAGGAGAGAGAGAAAGAAATTCAGTTTATAGTATGTCCTC 1080
DB 1021 TTCTCTAGATTAGTATAAAGGAGAGAGAGAAAGAAATTCAGTTTATAGTATGTCCTC 1080
QY 1081 TGAGATGTGCGGTTTGGCAACATAGCCACCGTAAATCATATAGCTCATATAGTGCCTACGTCA 1140
DB 1081 TGAGATGTGCGGTTTGGCAACATAGCCACCGTAAATCATATAGCTCATATAGTGCCTACGTCA 1140
QY 1141 GGTGGGCGAGCTCTGTGTCATCTCATATGATGATACATATGCTTTGTTCAACCGTTGTC 1200
DB 1141 GGTGGGCGAGCTCTGTGTCATCTCATATGATGATACATATGCTTTGTTCAACCGTTGTC 1200
QY 1201 TTGTTTCATGTCGCAAGCCCTTGCCTATTTCTGTAACCAAGGATATCTACTCCCAACAT 1260
DB 1201 TTGTTTCATGTCGCAAGCCCTTGCCTATTTCTGTAACCAAGGATATCTACTCCCAACAT 1260
QY 1261 CCATCTTACTCATGCACTTCCATGCAAAACAGCACATATGTTTCTGTAACCAATCAT 1320
DB 1261 CCATCTTACTCATGCACTTCCATGCAAAACAGCACATATGTTTCTGTAACCAATCAT 1320
QY 1321 AAAGATCACAACAGCTAGCGTTCTCCGCTAGCTTCCCTCTCTCTGCGCATCTTTT 1380
DB 1321 AAAGATCACAACAGCTAGCGTTCTCCGCTAGCTTCCCTCTCTCTGCGCATCTTTT 1380
QY 1381 CGTCCACGACCATG 1394
DB 1381 CGTCCACGACCATG 1394

RESULT 2
AAH76332
ID AAH76332 standard, DNA; 1394 BP.
AC AAH76332;
AC
AC
DT 29-OCT-2001 (first entry)
XX
XX
DE 2. mays Ms45 male tissue-preferred regulatory region encoding DNA.
XX
XX Ms45; male tissue; regulatory region; transcription; male fertility;
KM hybrid seed; ds.
XX
XX Zea mays.
OS
PN WO200160997-A2.
XX
XX 23-AUG-2001.
PD
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
XX (PION-) PIONEER HI-BRED INT INC.
PA
PI Albrechtsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
XX WPI; 2001-514772/56.
DR
XX
XX A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the Ms45 gene useful for
PT mediating fertility in a male plant.
XX
XX
PS Claim 4; Page 46; 50pp; English.
XX
XX The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the Ms45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The

CC	present sequence represents a nucleic acid sequence encoding an M45 male
CC	-tissue preferred regulatory region from Z. mays
XX	Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;
5Q	

Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;

Query Match	100.0%	Score 1394;	DB 5;	length 1394;
Best Local Similarity	100.0%	Pred. No. 5.1e-291;		
Matches 1394: Conservative	0	Mismatches	0	Indels 0

OY	1	CCATGSGTCTCTATGAAAAAGAGAGTCAATGTCCTATATACGTTTTCTTAAGGTC	60
Db	1	CCATGSGTCTCTATGAAAAAGAGTCAATGTCCTATATACGTTTTCTTAAGGTC	60
OY	61	CTTCTTCGCTTAATTACTGATCGAATCGGGGTACAAAAAATTCCACGSGTGATGAT	120
Db	61	CTTCTTCGCTTAATTACTGATCGAATCGGGGTACAAAAAATTCCACGSGTGATGAT	120
OY	121	CTCCATGTTCCACTTCTCCACCTCGCGTTGCAATTTCTTGAGTGCGGTGGTTCCAT	180
Db	121	CTCCATGTTCCACTTCTCCACCTCGCGTTGCAATTTCTTGAGTGCGGTGGTTCCAT	180
OY	181	CTGACCGAGGGCCATTCAGACACCTTTGGGAGACCCATCAAGGCGCTTTCGATGCGCC	240
Db	181	CTGACCGAGGGCCATTCAGACACCTTTGGGAGACCCATCAAGGCGCTTTCGATGCGCC	240
OY	241	CGAGACGATCGGGTCTGTGATCCAGGGGATATATGTCCCCCACAATCGTACCTATA	300
Db	241	CGAGACGATCGGGTCTGTGATCCAGGGGATATATGTCCCCCACAATCGTACCTATA	300
OY	301	TTATTATTCTTATAGATATTATTAAATTTTGGAAAAATAACAACTTATCTTTGTGTA	360
Db	301	TTATTATTCTTATAGATATTATTAAATTTTGGAAAAATAACAACTTATCTTTGTGTA	360
OY	361	GGGCGCTCAGCATAGATTTTTCGCTTAGGGGCCAGAAATGCGAGGACCAAGCATGTCTAGTG	420
Db	361	GGGCGCTCAGCATAGATTTTTCGCTTAGGGGCCAGAAATGCGAGGACCAAGCATGTCTAGTG	420
OY	421	TCCACTATTGGCAGTACCAGAAACAGATTTTAAAAAATATACCAAGTAACTAATCCACT	480
Db	421	TCCACTATTGGCAGTACCAGAAACAGATTTTAAAAAATATACCAAGTAACTAATCCACT	480
OY	481	CGAAAGCTATCATGTATGTTTAAAGAAACATCTATTAACCCAGATCCTCTTAAAAA	540
Db	481	CGAAAGCTATCATGTATGTTTAAAGAAACATCTATTAACCCAGATCCTCTTAAAAA	540
OY	541	CAGCATATTTGAAAAGAGACAAATATGTTCAGTTTCAAAACATCTAAGAGCGACAA	600
Db	541	CAGCATATTTGAAAAGAGACAAATATGTTCAGTTTCAAAACATCTAAGAGCGACAA	600
OY	601	TTATATCGAAGGTAAAGCTATAGCGTTTCTTTTCATCTTGTAATTTGTT	660
Db	601	TTATATCGAAGGTAAAGCTATAGCGTTTCTTTTCATCTTGTAATTTGTT	660
OY	661	ATTGTTTTTATATACATTTCTTCTCTTACAATAGAGTATTTCTTCGATTTTAA	720
Db	661	ATTGTTTTTATATACATTTCTTCTCTTACAATAGAGTATTTCTTCGATTTTAA	720
OY	721	ATGACTATTAAGTCAATTTTATATAAGACAGCATGTCGTAGATTTCTGTTCAAAATC	780
Db	721	ATGACTATTAAGTCAATTTTATATAAGACAGCATGTCGTAGATTTCTGTTCAAAATC	780
OY	781	TTTCTGATTTTTTTAAGAGCTAGTTGGACCCGTCTTCTTCAAGAAATTTGATTT	840
Db	781	TTTCTGATTTTTTTAAGAGCTAGTTGGACCCGTCTTCTTCAAGAAATTTGATTT	840
OY	841	TTCAAAAAAATTAGTTTATTTTCTCTTTATAAATAGAAAACATTAGAAAAATAGAGT	900
Db	841	TTCAAAAAAATTAGTTTATTTTCTCTTTATAAATAGAAAACATTAGAAAAATAGAGT	900
OY	901	TGCCGAGCTAGCCCTTAGAATGTTTTTCCCAATAATTACAATCACTGTGTATATATTTG	960
Db	901	TGCCGAGCTAGCCCTTAGAATGTTTTTCCCAATAATTACAATCACTGTGTATATATTTG	960

QY	961	GCCAGCCCCCAATAATTATTTTAAACCGAAACTGAANAATGAGCGAAACCAAAATCTGAGCTAT	1020
Db	961	GCCAGCCCCCAATAATTATTTTAAACCGAAACTGAANAATGAGCGAAACCAAAATCTGAGCTAT	1020
QY	1021	TTCTCTGATTAAGTAAAAAGGAGAGAGAGAGAAATCAAGTTTAAGCATATGGTCCC	1080
Db	1021	TTCTCTGATTAAGTAAAAAGGAGAGAGAGAGAAATCAAGTTTAAGCATATGGTCCC	1080
QY	1081	TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATATAGCTATAGTGCCCTACGTCA	1140
Db	1081	TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATATAGCTATAGTGCCCTACGTCA	1140
QY	1141	GGTTCCGAGACTCTTCGTGTCACTCAATAGGCAATCACTAGCTTTGTTCAACGTTTGTC	1200
Db	1141	GGTTCCGAGACTCTTCGTGTCACTCAATAGGCAATCACTAGCTTTGTTCAACGTTTGTC	1200
QY	1201	TTGTTTCATCGTCCAAACGCTTTCCTATTCTGAACCAAGAGATTAACCTACCTCCAAACAAT	1260
Db	1201	TTGTTTCATCGTCCAAACGCTTTCCTATTCTGAACCAAGAGATTAACCTACCTCCAAACAAT	1260
QY	1261	CCATCTTAATCTATGCAACTTTCATATGCAAAACACGACATATGTTTCTGAAACCAATCCATT	1320
Db	1261	CCATCTTAATCTATGCAACTTTCATATGCAAAACACGACATATGTTTCTGAAACCAATCCATT	1320
QY	1321	AAAGATCAACAACGTAAGCGTTCTCCGCGTAGTCTTCCTCTCTGCGCGATCTTTT	1380
Db	1321	AAAGATCAACAACGTAAGCGTTCTCCGCGTAGTCTTCCTCTCTGCGCGATCTTTT	1380
QY	1381	CGTCCACCACCATG 1394	
Db	1381	CGTCCACCACCATG 1394	

RESULT 3
 AAX07409
 ID AAX07409 standard; DNA; 1394 BP.
 XX
 AC AAX07409;
 XX
 DT 08-JUN-1999 (first entry)
 XX
 DE Zea mays Ms45 male tissue-preferred regulatory region.
 XX
 KW Ms45; male; tissue-preferred; regulatory region; plant cells;
 KM plant tissue; differentiated; hybrid seed; fertility; ss.
 XX
 OS Zea mays.
 XX
 PN MO9859061-A1.
 XX
 PD 30-DEC-1998.
 XX
 PF 19-JUN-1998; 98WO-US012895.
 XX
 PR 23-JUN-1997; 97US-00880499.
 XX
 PA (PION-) PIONEER HI-BRED INT INC.
 XX
 PI Albertsen MC, Fox TW, Garnaat CW, Huffman GA, Kendall TL;
 XX
 DR WPI; 1999-105628/09.
 XX
 PT New nucleic acid encoding a Ms45 male tissue-preferred regulatory region
 XX
 PS - useful in mediating plant fertility, especially hybrid seed production.
 XX
 Claim 3; Page 23-24; 39pp; English.
 XX
 The sequence is that encoding an Ms45 male tissue-preferred regulatory
 CC region. It may be used in the construction of a vector for a method of
 CC producing exogenous genes in a male tissue- preferred manner, which is
 CC useful in restoring or conferring fertility, such as in hybrid seed
 CC production. In conferring fertility, a monocot/dicot plant is transformed
 CC with the exogenous nucleotide sequence (a male sterility gene, preferably

CC Ms45), which encodes a product selected from auxins, rolB and diptheria
 CC toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
 CC and infertile plants
 CC XX

SQ Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;

Query Match 99.7%; Score 1389.2; DB 2; Length 1394;
 Best Local Similarity 99.8%; Pred. No. 5.6e-290;
 Matches 1391; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 1 CCAAGTGTCTCTATGAAAAAGATGATACATGTGTCTATATCCGTTTCTTAAAGGCTC 60
DB 1 CCAATGGTCTCTATGAAAAAGATGATACATGTGTCTATATCCGTTTCTTAAAGGCTC 60
QY 61 CTCTCTCTGCTTATTAATGATACGGAATCGGGGTACAAAAAATTCACACGGGTGCATGAT 120
DB 61 CTCTCTCTGCTTATTAATGATACGGAATCGGGGTACAAAAAATTCACACGGGTGCATGAT 120
QY 121 CTCATGTCTCAGTCTCTCCACCTCGGTGACATTTCTGGATGTGGGTGCTCCAT 180
DB 121 CTCATGTCTCAGTCTCTCCACCTCGGTGACATTTCTGGATGTGGGTGCTCCAT 180
QY 181 CTGACCGAGGCCCATCAGACACCTTTGCGGACACCCATCAAGGCTTTGCGATGAGCCCA 240
DB 181 CTGACCGAGGCCCATCAGACACCTTTGCGGACACCCATCAAGGCTTTGCGATGAGCCCA 240
QY 241 CGAGAGTATCGGGTGTGTGTGATCCAGGGGATATATGTCCCAAGATGCTACCTATA 300
DB 241 CGAGAGTATCGGGTGTGTGTGATCCAGGGGATATATGTCCCAAGATGCTACCTATA 300
QY 301 TTATTAATCTTTAGATATTAATTAATTTTGAAGAAAAATACAATTAATCTTTGTGTA 360
DB 301 TTATTAATCTTTAGATATTAATTAATTTTGAAGAAAAATACAATTAATCTTTGTGTA 360
QY 361 GGGCTTCAGATATGATTTTGGCTTAAAGGCCAGAAATCGAGAGACCAAGCCATCTAATG 420
DB 361 GGGCTTCAGATATGATTTTGGCTTAAAGGCCAGAAATCGAGAGACCAAGCCATCTAATG 420
QY 421 TCCACTATTGGGCTCCCGAGAACAGATTTAATAAATAATACCAAGTAACTAATCCACT 480
DB 421 TCCACTATTGGGCTCCCGAGAACAGATTTAATAAATAATACCAAGTAACTAATCCACT 480
QY 481 CGAAGCTATCATGTAATGTTTAAAGAACATCTATTTAAACACGATCCTCTTAAAAA 540
DB 481 CGAAGCTATCATGTAATGTTTAAAGAACATCTATTTAAACACGATCCTCTTAAAAA 540
QY 541 CAAAGCATATTTGCAAGAGACAAATTAATGTTTACAGTTTACAAATCTTAAAGAGCAAA 600
DB 541 CAAAGCATATTTGCAAGAGACAAATTAATGTTTACAGTTTACAAATCTTAAAGAGCAAA 600
QY 601 TTATATCGAAGAGTAAGCTATGACGTTACAGATTTTCTTTTCAATCTGTATTTGT 660
DB 601 TTATATCGAAGAGTAAGCTATGACGTTACAGATTTTCTTTTCAATCTGTATTTGT 660
QY 661 ATTGTTTTTATATACATTTTCTCTTACATAGATGATTTTCTTCGATTTTATAA 720
DB 661 ATTGTTTTTATATACATTTTCTCTTACATAGATGATTTTCTTCGATTTTATAA 720
QY 721 ATGACTATAAAGTCAATTTTATATAAGAGCAGCAATGTCGTAAATCTCGTTCAAAAAC 780
DB 721 ATGACTATAAAGTCAATTTTATATAAGAGCAGCAATGTCGTAAATCTCGTTCAAAAAC 780
QY 781 TTCTGATTTTTTAAAGAGTAGTTTGGCAACCTGTTCTTCAAGAAATTTGATTT 840
DB 781 TTCTGATTTTTTAAAGAGTAGTTTGGCAACCTGTTCTTCAAGAAATTTGATTT 840
QY 841 TTCAAAAAAATAGTTTATTTCTCTTATAAATAAGAAAAACATTAAGAAAAATGAGT 900
DB 841 TTCAAAAAAATAGTTTATTTCTCTTATAAATAAGAAAAACATTAAGAAAAATGAGT 900
QY 901 TGGCAGACTAGCCCTAGAAATGTTTTGCCAATAAATTAACATCACTGTATATAATTTTG 960
DB 901 TGGCAGACTAGCCCTAGAAATGTTTTGCCAATAAATTAACATCACTGTATATAATTTTG 960

```

```

QY 961 GCCAGCCCAATAATTTTAAACGAAACCTGAAATCGAGCGAAACCAATCTGAGCTAT 1020
DB 961 GCCAGCCCAATAATTTTAAACGAAACCTGAAATCGAGCGAAACCAATCTGAGCTAT 1020
QY 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAATCAGTTTAAAGTATTGCTCC 1080
DB 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAATCAGTTTAAAGTATTGCTCC 1080
QY 1081 TGAAGTGTGCGGTTTGGCAACATAGCCCAATCATAGCTCATPAGTGCTTACGTCA 1140
DB 1081 TGAAGTGTGCGGTTTGGCAACATAGCCCAATCATAGCTCATPAGTGCTTACGTCA 1140
QY 1141 GGTGTGGAGCTCTGTGTGATCTGACATGAGATCTGATAGTCTTGTCAACGTTGTC 1200
DB 1141 GGTGTGGAGCTCTGTGTGATCTGACATGAGATCTGATAGTCTTGTCAACGTTGTC 1200
QY 1201 TTGTTTCATTCGTCAAGCCTTGCTATTTCTGAACCAAGAGATACCTACCAACAAT 1260
DB 1201 TTGTTTCATTCGTCAAGCCTTGCTATTTCTGAACCAAGAGATACCTACCAACAAT 1260
QY 1261 CCATCTTACTCATGCACTTCCATGCAACACGCAATATGTTTCTGAACCAATTCATT 1320
DB 1261 CCATCTTACTCATGCACTTCCATGCAACACGCAATATGTTTCTGAACCAATTCATT 1320
QY 1321 AAAAGATCAACAGCTAGGCTTCTCCGCTAGCTTCCGCTCTGCTGCGGATCTTTT 1380
DB 1321 AAAAGATCAACAGCTAGGCTTCTCCGCTAGCTTCCGCTCTGCTGCGGATCTTTT 1380
QY 1381 CGTCCACGACCATG 1394
DB 1381 CGTCCACGACCATG 1394

```

RESULT 4
 AAH76333
 ID AAH76333 standard; DNA; 1394 BP.
 XX
 AC AAH76333;
 DT 29-OCT-2001 (first entry)
 XX
 DE 2. mays Ms45 male tissue-preferred regulatory region encoding DNA.
 XX
 KW Ms45; male tissue; regulatory region; transcription; male fertility;
 KW hybrid seed; ds.
 XX
 OS Zea mays.
 XX
 PN MO20016097-A2.
 XX
 PD 23-AUG-2001.
 XX
 PF 13-FEB-2001; 2001WO-US004527.
 XX
 PR 15-FEB-2000; 2000US-0050487.
 XX
 PA (PION-) PIONEER HI-BRED INT INC.
 XX
 PI Albrechtsen MC, Fox TW, Garmaat CM, Huffman G, Kendall TL;
 XX
 DR WPI; 2001-514772/56.
 XX
 PT A male tissue-preferred regulatory region comprising nucleotide sequences
 PT essential for initiating transcription of the MS45 gene useful for
 PT mediating fertility in a male plant.
 XX
 PS Claim 4; Page 47; 50pp; English.
 XX
 CC The invention provides a male tissue-preferred regulatory region (1)
 CC comprising nucleotide sequences essential for initiating transcription of
 CC the MS45 gene. A method of mediating male fertility in a plant is
 CC provided that involves introducing an expression vector comprising a

CC promoter operably linked to (I) into a plant where the exogenous gene
 CC impacts male fertility of the plant and (II) controls expression of the
 CC exogenous gene. A method of producing hybrid seeds is also provided. The
 CC present sequence represents a nucleic acid sequence encoding an Ms45 male
 CC -tissue preferred regulatory region from Z. mays

XX Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;

Query Match 99.7%; Score 1389.2; DB 5; Length 1394;

Best Local Similarity 99.8%; Pred. No. 5.6e-290;

Matches 1391; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

Qy 1 CCAATGCTCTATGAAAAAGAGTACAAAGTGTCTATATCCCTTTCTTAAGGTC 60
Db 1 CCAATGCTCTATGAAAAAGAGTACAAAGTGTCTATATCCCTTTCTTAAGGTC 60
Qy 61 CTCTCTGCTTATTTACTGATCAATCGGGTTTACAAAAAATTCCACGGGTGCATGAT 120
Db 61 CTCTCTGCTTATTTACTGATCAATCGGGTTTACAAAAAATTCCACGGGTGCATGAT 120
Qy 121 CTCATGTTCCACTTCTCCACCTCGGGTTCACATTTCTTGATGTGGTGGTCCAT 180
Db 121 CTCATGTTCCACTTCTCCACCTCGGGTTCACATTTCTTGATGTGGTGGTCCAT 180
Qy 181 CTGACCGAGGCCCTTCAGACACCTTTGGGACACCCATCAAGGGCTTTGGATGGCCA 240
Db 181 CTGACCGAGGCCCTTCAGACACCTTTGGGACACCCATCAAGGGCTTTGGATGGCCA 240
Qy 241 CGAAGACGTATCGGGTGTGGTATCGAGGGATATATGTCCCCCAAAATGTCACCTATA 300
Db 241 CGAAGACGTATCGGGTGTGGTATCGAGGGATATATGTCCCCCAAAATGTCACCTATA 300
Qy 301 TTATATTTCTTATAGATATTTATTTTGGAAAAATTAACAACTTATCTTTGTGTA 360
Db 301 TTATATTTCTTATAGATATTTATTTTGGAAAAATTAACAACTTATCTTTGTGTA 360
Qy 361 GGGCTTCAGATGATTTTGGCTTAAGGCCCAAGAAATGCGAGACCGCATGTCTAGTG 420
Db 361 GGGCTTCAGATGATTTTGGCTTAAGGCCCAAGAAATGCGAGACCGCATGTCTAGTG 420
Qy 421 TCCATTTTGGCACTACCCAGAACAAAGATTTAAAAAAATTAACAAAGTAACTCACT 480
Db 421 TCCATTTTGGCACTACCCAGAACAAAGATTTAAAAAAATTAACAAAGTAACTCACT 480
Qy 481 CGAAAGCTATCATGTATGTTTAAAGAACATCTATTAACACGATCCTCTTAAAAAA 540
Db 481 CGAAAGCTATCATGTATGTTTAAAGAACATCTATTAACACGATCCTCTTAAAAAA 540
Qy 541 CAAGCATTTTCCAAAAGACAAATTTATTTACGTTTACAAACATTTAAGCGACAAA 600
Db 541 CAAGCATTTTCCAAAAGACAAATTTATTTACGTTTACAAACATTTAAGCGACAAA 600
Qy 601 TTAATATGAAAGGTAGCTATGACGTTCAATTTTCTTTTCAATTTCTTTATTTTGT 660
Db 601 TTAATATGAAAGGTAGCTATGACGTTCAATTTTCTTTTCAATTTCTTTATTTTGT 660
Qy 661 ATGTGTTTATATCATTTTCTTCTTACAAATAGATGATTTTCTTCGATTTTATAA 720
Db 661 ATGTGTTTATATCATTTTCTTCTTACAAATAGATGATTTTCTTCGATTTTATAA 720
Qy 721 ATGACTTAAAGTCATTTTATATAGACGCGATGTGTAATTTCTGTTCAAAATC 780
Db 721 ATGACTTAAAGTCATTTTATATAGACGCGATGTGTAATTTCTGTTCAAAATC 780
Qy 781 TTCTGATTTTAAAGCTAGTTTGGCAACCTGTTCTTTCAAAAGATTTTGTATTT 840
Db 781 TTCTGATTTTAAAGCTAGTTTGGCAACCTGTTCTTTCAAAAGATTTTGTATTT 840
Qy 841 TTCAAAAAATATGTTATTTTCTTTTATAAATAGAAAACATTAAGAAAATAGAGT 900
Db 841 TTCAAAAAATATGTTATTTTCTTTTATAAATAGAAAACATTAAGAAAATAGAGT 900
Qy 901 TGCACAGCTAGCCTTAATGTTTCCCAATAAATTCATCACTGTGTAATATATTTG 960

```

```

Db 901 TGCACAGCTAGCCTTAATGTTTCCCAATAAATTCATCACTGTGTAATATATTTG 960
Qy 961 GCCAGCCCATTAATATTTTAAACCGAACTGAATCGAGCGAAACCAATCTGACTAT 1020
Db 961 GCCAGCCCATTAATATTTTAAACCGAACTGAATCGAGCGAAACCAATCTGACTAT 1020
Qy 1021 TTCTTAGATTAAGTAAAAAGGAGAGAGAGAGAAATCACTTTTAAGTCACTGTCCC 1080
Db 1021 TTCTTAGATTAAGTAAAAAGGAGAGAGAGAGAAATCACTTTTAAGTCACTGTCCC 1080
Qy 1081 TGAGATGTGGGTTTGGCAACGATAGCCACCGTAATCAAGGCTATAGGCTAGCTCA 1140
Db 1081 TGAGATGTGGGTTTGGCAACGATAGCCACCGTAATCAAGGCTATAGGCTAGCTCA 1140
Qy 1141 GGTTCGGCAGCTTCGTGTATCTCATATGACATGACATGATGCTTTCAACGGTCTGC 1200
Db 1141 GGTTCGGCAGCTTCGTGTATCTCATATGACATGACATGATGCTTTCAACGGTCTGC 1200
Qy 1201 TTGTTTCATGCTCAAGCCTTGCTTATTTGAAACCAAGAGATACCTACTCCAAACAT 1260
Db 1201 TTGTTTCATGCTCAAGCCTTGCTTATTTGAAACCAAGAGATACCTACTCCAAACAT 1260
Qy 1261 CCATCTTACTCATGCAACTTCCATGCAACAGCAATATGTTTCTGTAACCAATTCATT 1320
Db 1261 CCATCTTACTCATGCAACTTCCATGCAACAGCAATATGTTTCTGTAACCAATTCATT 1320
Qy 1321 AAAGATCAACAAGCTAGAGGTTCTCCGCTAGCTTCCCTCTCTCTGCGCATCTTTT 1380
Db 1321 AAAGATCAACAAGCTAGAGGTTCTCCGCTAGCTTCCCTCTCTCTGCGCATCTTTT 1380
Qy 1381 CGTCACACACCATG 1394
Db 1381 CGTCACACACCATG 1394

```

RESULT 5

AAH76340

ID AAH76340 standard; DNA; 255 BP.

AC AAH76340;

DT 29-OCT-2001 (first entry)

XX Z. mays Ms45 promoter fragment.

XX Ms45; male tissue; regulatory region; transcription; male fertility;

XX hybrid seed; promoter; de.

XX Zea mays.

XX WO200160997-A2.

XX 23-AUG-2001.

XX 13-FEB-2001; 2001WO-US004527.

XX 15-FEB-2000; 2000US-00504487.

XX (PION-) PIONEER HI-BRED INT INC.

XX Albetzen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;

XX WPI; 2001-514772/56.

XX A male tissue-preferred regulatory region comprising nucleotide sequences

XX essential for initiating transcription of the Ms45 gene useful for

XX mediating fertility in a male plant.

XX Example 5; Fig 8; 50pp; English.

CC The invention provides a male tissue-preferred regulatory region (I)

CC comprising nucleotide sequences essential for initiating transcription of

CC the MS45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a Z. mays Ms45 promoter fragment

XX Sequence 255 BP; 59 A; 86 C; 39 G; 71 T; 0 U; 0 Other;

Query Match 17.1%; Score 238.8; DB 5; Length 255;
Best Local Similarity 97.2%; Pred. No. 7.5e-42;
Matches 243; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1145 CGGAGCTGTCGTGTCATCCATGCGATACATGCTGTTGCAACGGTGTCTGTG 1204
DB 5 CCGGGATCCCGGTGTCATCCATGCGATACATGCTGTTGCAACGGTGTCTGTG 64
QY 1205 TCCATGCTCCAGCTGCTTCCATTTCTGAAACCAAGAGATACCTACTCCCAACATCCAT 1264
DB 65 TCCATGCTCCAGCTGCTTCCATTTCTGAAACCAAGAGATACCTACTCCCAACATCCAT 124
QY 1265 CTTACTGATGCACTTTCATGCAAAACGCAATATGTTTCTGAAACCAATCCATTAAAG 1324
DB 125 CTTACTGATGCACTTTCATGCAAAACGCAATATGTTTCTGAAACCAATCCATTAAAG 184
QY 1325 ATCAACAACAGCTAGCGTTTCTCCGCTAGCTTCCCTCTGCGCGATCTTTTCGTC 1384
DB 185 ATCAACAACAGCTAGCGTTTCTCCGCTAGCTTCCCTCTGCGCGATCTTTTCGTC 244
QY 1385 CACCACCATTG 1394
DB 245 CACCACCATTG 254

RESULT 6
AAH76334
ID AAH76334 standard; DNA; 158 BP.

XX AAH76334;

DT 29-OCT-2001 (first entry)

XX Z. mays Ms45 male tissue-preferred regulatory region fragment.

KM Ms45; male tissue; regulatory region; transcription; male fertility;

XX hybrid seed; ds.

XX Zea mays.

XX WO200160997-A2.

XX 23-AUG-2001.

PF 13-FEB-2001; 2001WO-US004527.

XX 15-FEB-2000; 2000US-00504487.

XX (PION-) PIONEER HI-BRED INT INC.

PI Albrechtsen MC, Fox TW, Garmaat CW, Huffman G, Kendall TL;

DR WPI; 2001-514772/56.

PT A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the MS45 gene useful for
XX mediating fertility in a male plant.

PS Claim 5; Page 47; 50pp; English.

XX The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the MS45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a

CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a DNA fragment -38 to -195 bases upstream of
CC the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
XX nucleotide sequence

XX Sequence 158 BP; 41 A; 50 C; 21 G; 46 T; 0 U; 0 Other;

Query Match 10.5%; Score 146; DB 5; Length 158;
Best Local Similarity 99.4%; Pred. No. 7.1e-22;
Matches 157; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1155 CGTGCATCTCCATGCGATGCTACTGTTTCAACCGTTCGTC-TTGTCCATCGTC 1213
DB 1 CGTGCATCTCCATGCGATGCTACTGTTTCAACCGTTCGTTGTTCCATCGTC 60
QY 1214 CAAGCTTGCCTATTCTGAAACCAAGAGATACCTACTCCCAACATCCATCTACTCAT 1273
DB 61 CAAGCTTGCCTATTCTGAAACCAAGAGATACCTACTCCCAACATCCATCTACTCAT 120
QY 1274 GCAACTTCCATGCAAAACGCAATATGTTTCTGAAAC 1311
DB 121 GCAACTTCCATGCAAAACGCAATATGTTTCTGAAAC 158

RESULT 7
AAL15210/C
ID AAL15210 standard; cDNA; 883 BP.

XX AAL15210;

DT 07-DEC-2001 (first entry)

XX Human breast cancer expressed polynucleotide 7667.

XX Human; breast cancer; cell marker; cytostatic; ss.

XX Homo sapiens.

XX WO200151628-A2.

PD 19-JUL-2001.

PF 10-JAN-2001; 2001WO-US0000798.

PR 14-JAN-2000; 2000US-0176077P.

PR 14-MAR-2000; 2000US-0189167P.

PR 24-MAR-2000; 2000US-0192099P.

PR 29-MAR-2000; 2000US-0193480P.

PR 15-MAY-2000; 2000US-0205230P.

PR 09-JUN-2000; 2000US-0211315P.

PR 25-JUL-2000; 2000US-0220534P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Lillie J, Xu Y, Wang Y, Steinmann K;

DR WPI; 2001-451856/48.

PT New peptide useful as a marker for the diagnosis of breast cancer.

PS Claim 1; Page 1378; 3695pp; English.

XX The invention relates to human breast cancer expressed polynucleotides
CC (AAL07544-AAL26789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterizing, treating and encoded
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity

XX Sequence 883 BP; 322 A; 18 C; 23 G; 396 T; 0 U; 124 Other;
SQ
Query Match 5.0%; Score 70.2; DB 4; Length 883;
Best Local Similarity 39.2%; Pred. No. 2.6e-05;
Matches 237; Conservative 0; Mismatches 367; Indels 0; Gaps 0;
QY 419 TGTCCACTATTGGCAGTACCCAGAACAGATTAAAAAAATTAACCAAGTAAGTAATCCA 478
DB 873 TGTNNNNNNNACCTNACCTTTAAANATNNNANTNNANNNAAAAATATTTAANN 814
QY 479 CTCGAAGCTATCATGTATGTTTAAAGAAACATCTATTAACCGACGATCTCTTAA 538
DB 813 TAAAAATTTTNNAAAAAATAATTTTANNTANNTANNTATTTTNTAAAAAANNAAAA 754
QY 539 AACAGCATATTTTGAAGAGACAAATATGTACAGTTTACAAACATCTAGAGCGACA 598
DB 753 TTAANNNTTTNTAANTTATTAACCAAAATTTTAAAAAATTTTAAAAANTT 694
QY 599 AATTATATCGAAAGGTAGCTAGAGCTGAGATTTTCTTTTCATTTCTGTATTG 658
DB 693 AATTAATTTNAATTTTNTAATNAAAAAATTTTNAATTTTANCAANTNTTTTTT 634
QY 659 TTAATGTTTTTATATACATTTCTCTCTACATAGAGATTTTCTCGATTTTATA 718
DB 633 TTTNTTTNNATTAATAAATTTTAAATTAATAAANNANNTTTTNTANNAATATTTA 574
QY 719 AATGACTATTAAGCTATTTTATATAGAGACGATGCTAGATTCGTTCAAAA 778
DB 573 AAAAAAANNTTTTNTTATANTTTATATAAATAATTTATNTNTTTTCTNAAAA 514
QY 779 TCTTTCTGATTTTTTAAGAGCTAGTTGGCAACCGTCTTCTTCAAGATTTTGATT 838
DB 513 AAAAAAATTTTNNAAANTTTTAAACCTTNAATTAANANANATTTTNTNTT 454
QY 839 TTTTCAAAAAATTTAGTTTATTTTCTCTTATATAAATAGAAAACTTAGAAAAATAG 898
DB 453 NNNANNTTAAAAATTTATTTTNTTANTTTATTAATTAANANANANANATTAAT 394
QY 899 GTTGCCAGACTAGCCCTAGAAATGTTTCCCATTAATTAACATCACTGTGTATATAT 958
DB 393 TTTNAAAAAATAATTTATANTTATNTANNAANANAAATTAANANANAAAAATTTTA 334
QY 959 TGGCCAGCCCATTAATTTATTAACCGAAATCGAAATCGAGCAACCAATCTGAGCT 1018
DB 333 NTTTAATTAATTAATAATTAANTTTAAACATCAATTNTTATTTTATTAATTAATAA 274
QY 1019 ATTT 1022
DB 273 NTTT 270
RESULT 8
ACN85231/c
ID ACN85231 standard; DNA; 960 BP.
ACN85231;
02-DEC-2004 (first entry)
Breast Cancer related marker, seq id 6181.
Cancer; breast; tumour; cytostatic; marker; detection; therapy; dr.
Homo sapiens.
US2003099974-A1.
29-MAY-2003.
18-JUL-2002; 2002US-00198846.
18-JUL-2001; 2001US-0306220P.

XX (MILL-) MILLENNIUM PHARM INC.
PA Lillie J, Xu Y, Wang Y, Steinmann K;
PI WPI; 2003-787014/74.
PT Novel isolated polypeptide associated with breast cancer, useful for
PT detecting presence of polypeptide in sample, as a marker for breast
PT cancer.
XX
PS Disclosure; SEQ ID NO 6381; 36pp; English.
CC The invention relates to an isolated polypeptide (I) associated with
CC breast cancer which is encoded by a nucleic acid molecule comprising a
CC nucleotide sequence (S1). Further disclosed is an antibody that binds to
CC the polypeptide of the invention. The activity of the polypeptide of the
CC invention may be described as cytostatic. The antibody is useful for
CC detecting the presence of (I) in a sample. Nucleic acid molecules of the
CC invention are useful in the detection of breast tumours. (I) is useful as
CC a marker for breast cancer and in breast cancer therapy. Sequences given
CC in records ACN78851-ACN92934 represent nucleic acid markers associated
CC with breast cancer. Note: The sequence listing does not form part of the
CC specification but may be obtained in electronic format from the USPTO web
CC site at seqdata.uspto.gov/sequence.html?docid=2003099974
XX
SQ Sequence 960 BP; 340 A; 33 C; 39 G; 421 T; 0 U; 127 Other;
Query Match 5.0%; Score 70.2; DB 11; Length 960;
Best Local Similarity 39.2%; Pred. No. 2.6e-05;
Matches 237; Conservative 0; Mismatches 367; Indels 0; Gaps 0;
QY 419 TGTCCACTATTTGGCAGTACCCAGAACAGATTAAAAAAATTAACCAAGTAATTAATCCA 478
DB 933 TGTNNNNNNNACCTNACCTTTAAANATNNNANTNNANNNAAAAATATTTAANN 874
QY 479 CTCGAAGCTATCATGTATGTTTAAGAAACATCATTAACCAACGATCTCTTAA 538
DB 873 TAAAAATTTTNNAAAAAATAATTTTANNTANNTAATTTTNTAAAAAANNAAAA 814
QY 539 AACAGCATATTTTGAAGAGACAAATATGTACAGTTTACAAACATCTAGAGCGACA 598
DB 813 TTAANNTTTTNTAANTTTATTAACCAAAATTTTAAAAAATTTTAAAAANTT 754
QY 599 AATTATATCGAAAGGTAGCTAGAGCTGAGATTTTCTTTTCATTTCTGTATTG 658
DB 753 AATTAATTTNAATTTTNTAATNAAAAAATTTTNAATTTTANCAANTNTTTTTT 694
QY 659 TTAATGTTTTTATATACATTTCTCTCTACATAGAGATTTTCTCGATTTTATA 718
DB 693 TTTNTTTNNATTAATAAATTTTAAATTAANANANANNTTTTNTAANANATTTTA 634
QY 719 AATGACTATTAAGCTATTTTATATAGAGACGATGCTAGATTCGTTCAAAA 778
DB 633 AAAAAAANNTTTTNTTATANTTTATATAAATAATTTTNTNTTTTCTNAAAA 574
QY 779 TCTTTCTGATTTTTTAAGAGCTAGTTGGCAACCGTCTTCTTCAAGAAATTTTGATT 838
DB 573 AAAAAAATTTTNNAAANTTTTAAACCTTNAATTAANANANATTTTNTNTT 514
QY 839 TTTTCAAAAAATTTAGTTTATTTTCTCTTATATAAATGAAAACTTAGAAAAATAGA 898
DB 513 NNNANNTTAAAAATTTATTTTNTTANTTTATTAATTTAANANANANATTAAT 454
QY 899 GTTGCCAGACTAGCCCTAGAAATGTTTCCCATTAATTAACATCACTGTGTATATAT 958
DB 453 TTTNAAAAAATAATTTATANTTATNTANNAANANAAATTAANANANAAAAATTTTA 394
QY 959 TGGCCAGCCCATTAATTTTAAACCGAAATCGAAATCGAGCAACCAATCTGAGCT 1018
DB 393 NTTTAATTAATTAATAATTAANTTTAAACATCAATTNTTATTTATTAATTAATAA 334
QY 1019 ATTT 1022

Db 333 NTTT 330

RESULT 9
ADRo4296
ID ADRo4296 standard; DNA; 13400 BP.

XX ADRo4296;

DT 04-NOV-2004 (first entry)

DE Corn FT homologue nucleotide sequence SEQ ID NO:63.

XX flowering locus T gene; FT; terminal flower; TFL; Apetala3, Ap3; plant;
KM floral development; plant sterility; plant fertility; flowering time;
KM plant growth rate; inflorescence architecture; tissue culture morphology;
KM cell division; FT homologue; gene; ds.

XX Zea mays.

XX WO2004067723-A2.

PD 12-AUG-2004.

PF 29-JAN-2004; 2004WO-US002422.

XX 30-JUN-2003; 2003US-00343477.

PA (PION-) PIONEER HI-BRED INT INC.
PA (DUPO) DU PONT DE NEMOURS & CO E I.

PI Danilevskaya O, Hermon P, Bruggemann E, Shibrproun D, Ananiev E;
PI Rafalski JA, Sakai H, Cahoon B, Cahoon R, Klein T;

DR WPI; 2004-580996/56.

PT New polynucleotides, specifically nucleic acid fragments encoding
PT flowering locus T gene (FT) or terminal flower (TFL), or Apetala3 (Ap3)
PT homologue, useful for floral development, e.g. engineering plant flowering
PT time.

PS Claim 6; SEQ ID NO 63; 109pp; English.

CC The present invention describes an isolated polynucleotide comprising a
CC first, second, third, fourth or fifth nucleotide sequence, or their
CC complement encoding a polypeptide either having flowering locus T gene
CC (FT), terminal flower (TFL), or Apetala3 (Ap3) homologue activity. Also
CC described: (1) a vector comprising the polynucleotide; (2) a recombinant
CC DNA construct comprising the polynucleotide; (3) transforming a cell by
CC transforming a cell with the polynucleotide; (4) a cell comprising the
CC recombinant DNA construct; (5) producing a plant comprising transforming
CC a plant cell with the polynucleotide, and regenerating a plant from the
CC transformed plant cell; (6) a plant comprising the recombinant DNA
CC construct; (7) a seed comprising the recombinant DNA construct; (8) an
CC isolated polynucleotide comprising a first nucleotide sequence, where the
CC first nucleotide sequence contains at least 30 nucleotides, and where the
CC first nucleotide sequence is comprised by another polynucleotide, where
CC the other polynucleotide includes the second, third, fourth, fifth or
CC sixth nucleotide sequence; (9) an isolated polypeptide having FT or Ap3
CC homologue activity, as described above; and (10) isolating a polypeptide
CC encoded by the polynucleotide comprising isolating the polypeptide from a
CC cell containing a recombinant DNA construct comprising the polynucleotide
CC operably linked to a regulatory sequence. The polynucleotides are useful
CC for floral development, e.g. engineering plant sterility/fertility,
CC flowering time, plant growth rate, inflorescence architecture, and tissue
CC culture morphology and the rate of cell division to enhance
CC transformation. The present sequence represents an FT homologue
CC nucleotide sequence from the present invention.

XX Sequence 13400 BP; 3962 A; 2948 C; 2695 G; 3795 T; 0 U; 0 Other;

Query Match 5.0%; Score 69.8; DB 13; Length 13400;

Best Local Similarity 73.6%; Pred. No. 6.3e-05;
Matches 89; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 811 ACCCTGTTCTTCAAGAAATTTGATTTTCAAAAAAATAGTTATTTCTCTTAA 870

DB 8559 ACCATTTTATTTTCAAGAGTTTATTTATCAAGAAAAATAGTTCATTTCTCTGG 8618

QY 871 TAAATAGAAAACCTTAGAAAATAGAGTTGCCAGACTAGCCCTAGAAATGTTCCCAA 930

DB 8619 AAAAATTAATAATCATTAGAAAAATGGGCTGTCAACTAGTCTTATTAGTTTCAT 8678

QY 931 T 931

DB 8679 T 8679

RESULT 10
AA210551
ID AA210551 standard; DNA; 2657 BP.

XX AA210551;

DT 16-NOV-1999 (first entry)

DE DNA sequence of the P-Zcp promoter of maize.

XX P-Zcp promoter; male sterile plant; glyphosate tolerance; glyphosate;
KM male reproductive tissue; hybrid seed production; crop outcrossing;
KM flower life; ss.

XX Zea mays.

PN WO9946396-A2.

PD 16-SEP-1999.

PF 09-MAR-1999; 99WO-US005126.

XX 09-MAR-1998; 98US-0077277P.

PA (MONS) MONSANTO CO.

PI Brown SM, Fromm ME;

DR WPI; 1999-551420/46.

PT Production of male sterile plants using a gene encoding glyphosate
PT tolerance, used for, e.g. production of hybrid seed.

PS Disclosure; Fig 1A-B; 54pp; English.

CC The present sequence represents the P-Zcp promoter of maize. It is used
CC in the method of the invention. The specification describes a method for
CC the production of male sterile plants. The method comprises selective
CC expression of DNA encoding a protein that causes tolerance to glyphosate
CC and application of glyphosate. The method uses two DNA molecules, each
CC operably linked to a separate promoter, whereby the first promoter
CC functions in plant cells to produce a protein that causes tolerance to
CC glyphosate, and the second promoter functions in plant cells to cause the
CC production of a second RNA sequence in a male reproductive tissue.
CC Expression of the DNA promotes tolerance to glyphosate in those tissues
CC in which it is expressed. Expression of the second DNA molecule causes
CC the production of an RNA sequence which can inhibit the glyphosate
CC tolerance generated by expression of the first DNA molecule. By using a
CC promoter for the second DNA molecule which restricts the production of
CC the antisense RNA to only a subset of the tissues which express the first
CC DNA molecule, only the subset of tissues in which the second DNA molecule
CC is expressed will be susceptible to glyphosate toxicity. In this way, a
CC specific cell type or combination of cell types, depending upon the
CC promoters utilized, can be selectively ablated by application of
CC glyphosate to the plant. The methods can be used for producing male-
CC sterile plants for use in the production of hybrid seed, for minimizing
CC undesirable crop outcrossing, and for lengthening flower life. The

CC methode can be used with plants such as corn, wheat, rice, canola, oat,
CC barley, alfalfa, carrot, cotton, oilseed, oilseed rape, sugarbeet,
CC sunflower, soybean, tomato, cucumber and squash
XX

SO Sequence 2657 BP; 686 A; 611 C; 677 G; 683 T; 0 U; 0 Other;

Query Match 4.8%; Score 67.2; DB 2; Length 2657;

Best Local Similarity 74.2%; Pred. No. 0.00015; Mismatches 33; Indels 1; Gaps 1;

Matches 98; Conservative 0; Mismatches 33; Indels 1; Gaps 1;

QY 787 ATTTTAAAGCTAGTTGGCAACCTGTTCTTCAAGATTGTTGTTTCAAA 846

DB 1883 ATCTTTTAGGGCTAGTTGGGAACCACTTT-TTCCAAGGATTTTCATTTTCCGAG 1941

QY 847 AAAAATTAGTTTCTTTTATATAAATGAAAAACATTAGAAAAATGAGTTGCCAG 906

DB 1942 GGAATAATGATTCATTTCCCTTGGAATAATATGCAATCCATGGAAAAATGTGTTCCAA 2001

QY 907 ACTAGCCCTAGA 918

DB 2002 ACTAGCCCTAAA 2013

RESULT 11

AAKS8751/c

ID AAKS8751 standard; cDNA; 6027 BP.

XX AAKS8751;

AC AAKS8751;

DT 16-AUG-1999 (first entry)

XX Maize d111 gene encoding starch synthase enzyme D11.

XX Starch synthase; SSI; D11; d111 gene; maize; transgenic plant; ss.

XX Zea mays.

XX Key

FT m18c_feature

FT m18c_feature

FT m18c_feature

FT m18c_feature

FT m18c_feature

FT m18c_feature

FT m18c_feature

FT m18c_feature

FT m18c_feature

FT m18c_feature

FT m18c_feature

FT m18c_feature

FT m18c_feature

XX WO924575-A1.

XX 20-MAY-1999.

PF 12-NOV-1998; 98WO-US024225.

XX 12-NOV-1997; 97US-00968542.

XX (IOWA) UNIV IOWA STATE RES FOUND INC.

XX Myers AM, James MG;

XX MPI; 1999-327406/27.

XX P-PSDB; AAY06199.

XX Nucleic acid encoding starch synthase enzyme of maize.

XX Claim 1; Page 104-107; 138pp; English.

XX This is the nucleotide sequence of the maize gene d111 (d11). To

XX illustrate the role of the d11 locus in starch biosynthesis, a transposon

XX tagging strategy was used to isolate the gene and describe its

XX polypeptide product. The invention reports tagging of the d11 locus with

XX Mu transposon, cloning and characterization of a portion of the gene, and

XX isolation of a near full-length cDNA (the present sequence). The amino acid

XX sequence (see AAY06199) deduced from this cDNA indicates that d11 codes

XX for a 186 kDa protein extremely similar to potato tuber starch synthase

XX SSI1. Its expression pattern indicates that d11 codes for SSI1 of maize

XX endosperm. The d11 product contains unique sequence features in its N-

XX terminus that may mediate direct interactions with other starch

XX biosynthetic enzymes. Mutations within the maize SSI1 gene affect

XX multiple aspects of starch biosynthesis by disrupting an enzyme complex

XX containing starch synthase(s), starch branching enzyme(s) and possibly

XX starch debranching enzyme(s). The isolated cDNA can be used to provide an

XX enzyme with which to regulate the production of starch, and with which to

XX produce altered or novel forms of starch, e.g. in transgenic plants.

XX Expression of d11 in bacteria and yeasts also modifies glycogen

XX production. Claimed expression vectors comprise the cDNA or fragments of

XX it that code for functional portions of d11

XX SO Sequence 6027 BP; 1800 A; 1177 C; 1441 G; 1609 T; 0 U; 0 Other;

XX Query Match 4.8%; Score 66.4; DB 2; Length 6027;

XX Best Local Similarity 77.5%; Pred. No. 0.00028;

XX Matches 93; Conservative 0; Mismatches 26; Indels 1; Gaps 1;

QY 792 TTTAAGAGCTAGTTGGCAACCTGTTCTTCAAGATTGTTGTTTCAAAAAA 851

DB 5606 TCTACGGGCTAGTTGGGAACCCCAATT-TTCCAAGGATTTTCATTTTCAAGAAAA 5548

QY 852 TTAGTTTATTTCTGTTTATAAATGAAAAACATTAGAAAAATGAGTTGCCAGCTAG 911

DB 5547 TTAGTTTATTTCTGTTTATAAATGAAAAACATTAGAAAAATGAGTTGCCAGCTAG 5488

RESULT 12

ABX09935/c

ID ABX09935 standard; DNA; 6027 BP.

XX ABX09935;

XX 17-FEB-2003 (first entry)

XX DNA encoding maize Starch synthase III (D11).

XX Starch; starch synthase; glucan association domain; GLASS; linker domain;

XX LINR; glucosyl transferase domain; GLYTR; C-terminal end; CTEND;

XX granule bound starch synthase; GBS; morphology; retrogradation;

XX waterbinding; swelling potential; gene; ds.

XX Zea mays.

XX WO200279410-A2.

XX 10-OCT-2002.

XX 29-MAR-2002; 2002WO-US009574.

XX	PR	30-MAR-2001; 2001US-0279720P.
XX	PA	(BADI) BASF PLANT SCI GMBH.
XX	PI	Commuri P, Keeling PL, Ramirez N, McKean A, Gao Z, Guan H;
XX	PT	WPI; 2003-040678/03.
XX	DR	New DNA encoding fusion protein consisting of 4 different functional domains selected from glucan association domain, linker domain, glucosyl transferase domain, and C-terminal end, useful for producing modified starches.
XX	PS	Claim 33; Page 225-227; 265pp; English.
XX	CC	The invention describes an isolated DNA molecule encoding a fusion protein consisting of 4 different functional domains selected from glucan association domain (GLAS), linker domain (LINKR), glucosyl transferase domain (GLYTR), and C-terminal end (CTEND) which are operably linked to one another. The DNA molecule is useful for expressing in plants polypeptides including starch synthase enzymes as fusion proteins with improved affinity to starch and modified catalytic capabilities and to the in vivo and in vitro synthesis of glucan chains of modified lengths as compared to plants producing native starch or starch produced with native starch synthases. Expression of the starch synthase fusion proteins along with granule bound starch synthase (GBSS) will lead to a modified starch having an altered or improved morphology, retrogradation, waterbinding, or swelling potential of the granules, gel strength, adhesiveness, cohesiveness, hardness, elasticity, increased or decreased granule size, degree of branching, crystallinity, degree of cross-linking, and increased or decreased glucan chain lengths. This sequence encodes a starch synthase used in the invention
XX	SQ	Sequence 6027 BP; 1800 A; 1177 C; 1441 G; 1609 T; 0 U; 0 Other;
YY	Query Match	4.8%; Score 66.4; DB 8; Length 6027;
YY	Best Local Similarity	77.5%; Pred. No. 0.00028;
YY	Matches 93; Conservative 0; Mismatches 26; Indels 1; Gaps 1;	
OY	792 TTAAAGACTGTGGGACCGCTTCTTTCAAGAATTTGATTTTTTCAAAAAAA 851	
Dd	5606 TCACGGGCTAGTTGGGAACCCCATTT-TTCCAAGGATTTCCATTTTTCAGAAAAA 5548	
OY	852 TTAGTTATTTCCTTTATAAATAGAAAAACATTAGAAAAATAGAGTGCACAGACTAG 911	
Dd	5547 TTAGTTATTTCCTTTATGGAATAATGAAATCTTTGAAAAAATAGAGTTCACTACTAG 5488	
RESULT 13		
ADK12106/C		
ID	ADK12106 standard; cDNA; 6027 BP.	
AC	ADK12106;	
XX	03-JUN-2004 (first entry)	
DE	cDNA encoding maize starch synthase, DU1.	
KM	Maize; duil 1; DU1; starch synthase;	
KM	alpha-1,4-glycosyltransferase catalytic activity; starch production;	
OS	Zea mays.	
XX	Key Location/Qualifiers	
FT	CDS /*tag= "a	
FT	/product= "DU1"	
DD	US2004049810-A1.	
DD	11-MAR-2004.	

```

XX 05-AUG-2003; 2003US-00634262.
PF 12-NOV-1997; 97US-00968467.
PR 12-NOV-2000; 2000US-00554467.
XX (MYER/) MYERS A. M.
PA (JAME/) JAMES M G.
XX Myers AM, James MG,
XX MPI: 2004-238526/22.
XX P-PSDB; ADK121117.
XX New nucleic acid designated dnull1, encoding a starch synthase, useful in
XX starch production.
XX Claim 1; SEQ ID NO 1; 58pp; English.
XX
XX The present invention relates to the isolation of a maize gene, dnull 1
XX (DUI), and the polypeptide it encodes. The DUI polypeptide has starch
XX synthase activity, and comprises an N-terminal arm region, a C-terminal
XX catalytic region, and a region of about 900 amino acids terminating with
XX the catalytic region. The C-terminal catalytic region has a catalytic
XX domain comprising alpha-1,4-glycosyltransferase catalytic activity. The
XX DUI polynucleotide sequence is useful in producing starch e.g. from a
XX transgenic plant or transfected cell. The present sequence encodes maize
XX DUI.
XX
XX Sequence 6027 BP; 1800 A; 1177 C; 1441 G; 1609 T; 0 U; 0 Other;
XX
XX Query Match 4.8%; Score 66.4; DB 12; Length 6027;
XX Best Local Similarity 77.5%; Pred. No. 0.00028;
XX Matches 93; Conservative 0; Mismatches 26; Indels 1; Gaps 1;
XX
XX 792 TTTAAGAGCTAGATTGGCAACCTGTTCTTTCAAGAATTTGATTTTCAAAAAA 851
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
XX 5606 TCTCTGCGGCTAGTTGGGAACCCCATTT-TTCCAAGGATTTCCATTTTCCAGAAAAA
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
XX 852 TTAGTTAATTTTCCTTATTAATAAGAAAAACCTTAGAAAAAATAGAGTGGCAGACTAG 911
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
XX 5547 TTAGTTAATTTTCATTGGAAAAAATTAATAATCTCTGGAAAAAATAGAGTTCACTACTAG 5488
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
XX
XX RESULT 14
XX ID ABX35844 standard; cDNA; 439 BP.
XX ABX35844
XX ABX35844;
XX
XX 20-FEB-2003 (first entry)
XX
XX Bovine EST associated with lactation/muscle/fat deposition #1009.
XX
XX Bovine; SB; EST; expressed sequence tag; lactation; LMFD;
XX muscle deposition; fat deposition; genome mapping; gene identification;
XX gene analysis; cattle breeding.
XX
XX Bos Taurus.
XX
XX US2002137139-A1.
XX
XX 26-SEP-2002.
XX
XX 24-SEP-2001; 2001US-00960352.
XX
XX 12-JAN-1999; 99US-0115707P.
XX
XX 11-JAN-2000; 2000US-00480902.
XX
XX (BYAT/) BYATT J C.
XX (MATH/) MATHIALAGAN N.
XX (TAON/) TAO N.
XX (WARR/) WARREN W C.

```

PI	Byatt JC, Machalaagan N, Tao N, Warren WC;
XX	WPI; 2003-110599/10.
DR	
XX	New nucleic acid associated with lactation, and muscle and fat
PT	deposition, useful for genome mapping, gene identification and analysis,
PT	cattle breeding, or for genetically improving cattle.
XX	
PS	Claim 2; SEQ ID NO 1009; 245bp; English.
XX	
CC	The invention relates to a purified nucleic acid molecule associated with
CC	lactation or muscle and fat deposition (designated LMP), derived from
CC	cattle, and the LMP nucleic acid can specifically hybridise to a second
CC	nucleic acid molecule comprising any of 15112 nucleotide sequences,
CC	appearing as ABX34836-ABX49947, or complements of them. Also included are
CC	; (1) a transformed cell having a nucleic acid comprising an LMP nucleic
CC	acid linked to a promoter and a 3' non-translated sequence that
CC	functions in the cell to cause termination of transcription and addition
CC	of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
CC	(2) determining a level or pattern of a molecule in a bovine cell or
CC	tissue comprising: (a) incubating a marker nucleic acid (comprising any
CC	of the 15112 nucleic acid sequences or its complement or fragment) with a
CC	complementary nucleic acid molecule obtained from the bovine cell or
CC	tissue, where hybridisation between the marker nucleic acid and the
CC	complementary nucleic acid permits the detection of the molecule; and (b)
CC	detecting the level or pattern of the complementary nucleic acid, where
CC	the detection of the complementary nucleic acid is predictive of the
CC	level or pattern of the molecule. The LMP nucleic acid is used for
CC	determining a level or pattern of a molecule in a bovine cell or tissue.
CC	It is useful for genome mapping, gene identification and analysis, cattle
CC	breeding, preparation of constructs for use in cattle gene expression, or
CC	for genetically improving cattle. The present sequence is one of the
CC	15112 bovine LMP EST (expressed sequence tag) nucleic acids. Note: The
CC	present sequence was not shown in the specification but was obtained in
CC	electronic format from the USPTO web site:
CC	seqdata.neptco.gov/Sequence.html?DocId=200201377139
XX	
SQ	Sequence 439 BP; 45 A; 51 C; 56 G; 286 T; 0 U; 1 Other;
	Query Match 4.7%; Score 66.2; DB 8; Length 439;
	Best local Similarity 54.3%; Pred. No. 0.00016;
	Matches 134; Conservative 0; Mismatches 113; Indels 0; Gaps 0
OY	626 TTCAAGATTTTCTTTTCATTCTGTTATTTGTATTTCTTTTAATACATTTCTCT
DB	182 TTCGTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTTATTTATTTTCTTT
OY	686 CTTCACAATAGACTGATTTTCTTCGAATTTATATAAAGCTATTTATATA
DB	242 TTTTATATTTTTTTTTTTTTTTATTTATTTTTTTTTTTTTTTTTTTTTTTT
OY	746 AGAGCAGCATGTCTGAGATTTCTGTTCAAATCTTCTGATTTTTPAAGCGTAGT
DB	302 TGTGGGGGGTGTT
OY	806 TGCGAACCTGTGTTCTTCAAGAATTTGATTTTTCAAAAAAATTAGTTATTTCT
DB	362 TTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTATTTATTTCTTTATTTTTTTTT
OY	866 CTTTATA 872
DB	422 TTCTTTA 428
RESULT 15	
AA187279/C	AA187279 standard; CDNA; 346 BP.
XX	AA187279;
AC	
TX	06-NOV-2001 (first entry)
DT	

```

DE Human polynucleotide SEQ ID NO 7339.
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX tissue growth factor; immunomodulatory; cancer; leukaemia;
XX nervous system disorders; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
XX WO200164835-A2.
XX
XX 07-SEP-2001.
XX
XX 26-FEB-2001; 2001WO-US004927.
XX
XX 28-FEB-2000; 2000US-00515126.
XX
XX 18-MAY-2000; 2000US-00577409.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX
XX WPI; 2001-514838/56.
XX
XX P-PSDB; AAO07348.
XX
XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
XX and treating e.g. leukemia, inflammation and immune disorders.
XX
XX Claim 1; SEQ ID NO 7339; 1399pp + Sequence Listing; English.
XX
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
XX the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and
XX activin/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer. Leukaemia, nervous system disorders, arthritis and
XX inflammation. Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic format
XX directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 346 BP; 188 A; 22 C; 16 G; 120 T; 0 U; 0 Other;

Query Match 4.7%; Score 65.6; DB 4; Length 346;
Best Local Similarity 52.6%; Pred. No. 0.0002;
Matches 143; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

QY 446 AGATTAAAAAATAACCAAGTACTAATCCACGACCGAAGCTATGCATGTAATGTTTAA 505
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 334 ACAAATTAAGATATATATTTAAATTAATAAGATTAAATTAATTAATCAATAAAAAATTTA 275
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 506 GAAACATCTATTAAAACCCAGATCCCTTTAAAAAACAAGCATATTTCCGAAAGAGCAAAAT 565
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 274 GACTTATATATATATTAATTAATTTTATTTTGAABAAATTAATTACATTTAAACAAATPAAAAA 215
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 566 TATGTACAGTTTACAAACATCTAAGACGCAAAATTAATTCGAAAGGTAAGCTATGACG 625
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 214 TTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTT 155
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 626 TTGAGATTTTCTTTTCATCTCTGTTATTTTGTTATTTGTTTATTAATTAATTAATTTCTCT 685
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 154 TATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 95
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 686 CTACAAATAGAGTATTTTCTCCGATTTTAT 717
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 94 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 63
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Search completed: September 15, 2005, 21:33:03
Job time : 679.892 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 14, 2005, 19:14:19 ; Search time 210.859 Seconds

(without alignments)
10817.505 Million cell updates/sec

Title: US-10-713-381-1

Perfect score: 1394
Sequence: 1 ccctggcgcctcctcgaataa.....cttttcgcaccaccaccag 1394

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A COMB. seq: *
2: /cgn2_6/ptodata/1/ina/5B COMB. seq: *
3: /cgn2_6/ptodata/1/ina/5C COMB. seq: *
4: /cgn2_6/ptodata/1/ina/5D COMB. seq: *
5: /cgn2_6/ptodata/1/ina/5E COMB. seq: *
6: /cgn2_6/ptodata/1/ina/5F COMB. seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1394	100.0	1394	3	US-08-880-499-1
2	1389.2	99.7	1394	3	US-08-880-499-2
3	66.4	4.8	6027	2	US-08-968-542C-1
4	66.4	4.8	6027	2	US-09-554-467A-1
5	58.6	4.2	2523	4	US-08-410-784A-3
6	57	4.1	1914	2	US-08-487-826B-13
7	55.8	4.0	279	4	US-09-313-294A-5397
8	55.6	4.0	612	4	US-09-902-540-1357
9	55.6	4.0	1394	3	US-08-880-499-1
10	55.6	4.0	1394	3	US-08-880-499-2
11	55.4	4.0	1055	4	US-09-806-708B-23
12	55.2	4.0	2614	4	US-09-004-056-1
13	55	3.9	336	4	US-09-640-173-53
14	55	3.9	336	4	US-09-713-550-53
15	55	3.9	336	4	US-09-825-294-53
16	55	3.9	336	4	US-09-970-866-53
17	54.6	3.9	5588	4	US-09-949-016-15129
18	53.8	3.9	307	4	US-09-313-294A-4743
19	53.8	3.9	6027	2	US-08-968-542C-1
20	53.8	3.9	6027	2	US-09-554-467A-1
21	52.2	3.7	1141	4	US-09-806-708B-22
22	51.8	3.7	1141	4	US-09-806-708B-22
23	51.6	3.7	134987	4	US-09-949-016-15348
24	51.6	3.7	134987	4	US-09-949-016-15349
25	51.6	3.7	134987	4	US-09-949-016-15350
26	51.6	3.7	134987	4	US-09-949-016-15507
27	51.6	3.7	134987	4	US-09-949-016-15508

28	51.6	3.7	134987	4	US-09-949-016-15509	Sequence 15509, A
29	51.4	3.7	50383	4	US-09-949-016-17600	Sequence 17600, A
30	51.4	3.7	129415	4	US-09-949-016-16997	Sequence 16997, A
31	51	3.7	601	4	US-09-949-016-156535	Sequence 156535, A
32	50.6	3.6	1039	4	US-09-902-540-1280	Sequence 1280, Ap
33	50.6	3.6	16573	4	US-09-949-016-14876	Sequence 14876, A
34	50.6	3.6	18773	4	US-09-949-016-14164	Sequence 14164, A
35	50.4	3.6	721	1	US-08-451-405A-2	Sequence 2, Appl1
36	50.2	3.6	832	4	US-09-621-976-2813	Sequence 2813, Ap
37	50.2	3.6	1392	4	US-09-949-016-16001	Sequence 16001, A
38	50	3.6	1392	4	US-09-257-584-1	Sequence 1, Appl1
39	49.8	3.6	2435	3	US-09-306-593-1	Sequence 1, Appl1
40	49.8	3.6	231129	4	US-09-949-016-16110	Sequence 16110, A
41	49.8	3.6	266283	4	US-09-949-016-11934	Sequence 11934, A
42	49.6	3.6	640681	4	US-09-790-988-1	Sequence 1, Appl1
43	49.2	3.5	126176	4	US-09-949-016-16137	Sequence 16137, A
44	49.2	3.5	126176	4	US-09-949-016-16138	Sequence 16138, A
45	47.8	3.4	601	4	US-09-949-016-25787	Sequence 25787, A

ALIGNMENTS

RESULT 1
US-08-880-499-1

Sequence 1, Application US/08880499
Patent No. 6037523

GENERAL INFORMATION:

APPLICANT: Albertson, Marc C.

APPLICANT: Fox, Tim W.

APPLICANT: Carl, Garraet W.

APPLICANT: Hufman, Gary A.

APPLICANT: Kendall, Timmy L.

TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.

STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.

STREET: Box 1000

CITY: Johnston

STATE: Iowa

COUNTRY: USA

ZIP: 50131

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/880,499

FILING DATE: CONCURRENTLY HERewith

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Sweeney, Patricia A.

REGISTRATION NUMBER: 32,733

REFERENCE/DOCKET NUMBER: 0578

TELEPHONE/COMMUNICATION INFORMATION:

TELEPHONE: (515) 248-4800

TELEFAX: (515) 248-4844

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1394 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-880-499-1

Query Match 100.0%; Score 1394; DB 3; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCATGGTGTCTATGAAAAAGATGAGTACATGTGTCTATCCGTTTCTTAAGGATCC 60
|
|
|
Db 1 CCATGGTGTCTATGAAAAAGATGAGTACATGTGTCTATCCGTTTCTTAAGGATCC 60
|
|
|
QY 61 CTTCCTTCTGCTTAATTACTGACTGAATCGGGGTTTACAAAAAATTCACCGGTCATGAT 120
|
|
|
Db 61 CTTCCTTCTGCTTAATTACTGACTGAATCGGGGTTTACAAAAAATTCACCGGTCATGAT 120
|
|
|
QY 121 CTCGATGTTCCATCTTCCCACTCGGTTGCACTTTCTTGATGTCCGTGTTCCCAT 180
|
|
|
Db 121 CTCGATGTTCCATCTTCCCACTCGGTTGCACTTTCTTGATGTCCGTGTTCCCAT 180
|
|
|
QY 181 CTGACCGAGGCCCATGAGACACCTTCGAGACACCCATCAAGGACCTTCGATGAGCCA 240
|
|
|
Db 181 CTGACCGAGGCCCATGAGACACCTTCGAGACACCCATCAAGGACCTTCGATGAGCCA 240
|
|
|
QY 241 CGAGAGGTATCGGGTCGTGTGATCCAGGGGATATATGTCGCCCAATGTCACCTATA 300
|
|
|
Db 241 CGAGAGGTATCGGGTCGTGTGATCCAGGGGATATATGTCGCCCAATGTCACCTATA 300
|
|
|
QY 301 TTATATCTCTTGAATTAATTAATTTTGGAAAAATTAACAATTAATCTTTGTGTA 360
|
|
|
Db 301 TTATATCTCTTGAATTAATTAATTTTGGAAAAATTAACAATTAATCTTTGTGTA 360
|
|
|
QY 361 GGGCTCAGCATGATTTTCCTTAAGGCCAGAAAAATGCGAGGACGAGCCATGCTAGTG 420
|
|
|
Db 361 GGGCTCAGCATGATTTTCCTTAAGGCCAGAAAAATGCGAGGACGAGCCATGCTAGTG 420
|
|
|
QY 421 TCCACTATTTGGCACTACCCAGAAACAAGATTTAAAAAATTAACCAAGTATCTATCCACT 480
|
|
|
Db 421 TCCACTATTTGGCACTACCCAGAAACAAGATTTAAAAAATTAACCAAGTATCTATCCACT 480
|
|
|
QY 481 CGAAGCTATCAGTATGTTTAAAGAAACATCTATTTAAACACGATCCTCTTAATAAAA 540
|
|
|
Db 481 CGAAGCTATCAGTATGTTTAAAGAAACATCTATTTAAACACGATCCTCTTAATAAAA 540
|
|
|
QY 541 CAAGCATATTTGAAAAAGACAATTAATGTTACAGTTTACCAACATCTAAGCGACAAA 600
|
|
|
Db 541 CAAGCATATTTGAAAAAGACAATTAATGTTACAGTTTACCAACATCTAAGCGACAAA 600
|
|
|
QY 601 TTATATCGAAAGGTAAAGCTATGACGTTTCAGATTTTCTTTTCATTTCTGTTATTTGTT 660
|
|
|
Db 601 TTATATCGAAAGGTAAAGCTATGACGTTTCAGATTTTCTTTTCATTTCTGTTATTTGTT 660
|
|
|
QY 661 ATTGTTTTATATACATTTTCTTCTTAATAGATGATTTTCTTCGATTTTATAA 720
|
|
|
Db 661 ATTGTTTTATATACATTTTCTTCTTAATAGATGATTTTCTTCGATTTTATAA 720
|
|
|
QY 721 ATGACTATAAAGTCATTTTATAAAGACGACATGCTAGATTCGTTCAAAAATC 780
|
|
|
Db 721 ATGACTATAAAGTCATTTTATAAAGACGACATGCTAGATTCGTTCAAAAATC 780
|
|
|
QY 781 TTTCTGATTTTAAAGAGCTAGTTGGCAACCTGTTTCTTCAAGAATTTGATTTT 840
|
|
|
Db 781 TTTCTGATTTTAAAGAGCTAGTTGGCAACCTGTTTCTTCAAGAATTTGATTTT 840
|
|
|
QY 841 TTCAAAAAAAATAGTTTATTTTCTCTTATAAATAGAAAAACATAGAAAAATAGAGT 900
|
|
|
Db 841 TTCAAAAAAAATAGTTTATTTTCTCTTATAAATAGAAAAACATAGAAAAATAGAGT 900
|
|
|
QY 901 TGGCAGACTAGCCCTAGAAATGTTTCCCAATAAATTAACATCACTGTATATTAATTG 960
|
|
|
Db 901 TGGCAGACTAGCCCTAGAAATGTTTCCCAATAAATTAACATCACTGTATATTAATTG 960
|
|
|
QY 961 GCGAGCCCCATAAATTTTAAACCGAAACTGAATCGAGCGAAACCAATCTGACTAT 1020
|
|
|
Db 961 GCGAGCCCCATAAATTTTAAACCGAAACTGAATCGAGCGAAACCAATCTGACTAT 1020
|
|
|
QY 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAATCAGTTTAAAGTCAATGCTCC 1080
|
|
|
Db 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAATCAGTTTAAAGTCAATGCTCC 1080
|
|
|
QY 1081 TGAGATGTGTGGGTTGGCAACGATAGCACCGTAAATCATAGTCTATAGTGTCTAGCTCA 1140
|
|
|

Db 1081 TGAGATGTGTGGGTTGGCAACGATAGCACCGTAAATCATAGTCTATAGTGTCTAGCTCA 1140
|
|
|
QY 1141 GGTGGGCAAGCTCTCTGTCTATCTCATGAGCATACTACCTGTTCAACCGTTGCTC 1200
|
|
|
Db 1141 GGTGGGCAAGCTCTCTGTCTATCTCATGAGCATACTACCTGTTCAACCGTTGCTC 1200
|
|
|
QY 1201 TTGTTCAATGTCGCAAGCCTTTCCTTCTGAAACCAAGGATTAATCTTCCCAACAT 1260
|
|
|
Db 1201 TTGTTCAATGTCGCAAGCCTTTCCTTCTGAAACCAAGGATTAATCTTCCCAACAT 1260
|
|
|
QY 1261 CCATCTTACTCATGCACTTCCATGCAACAGCAATATGTTTCTGAAACCAATCATT 1320
|
|
|
Db 1261 CCATCTTACTCATGCACTTCCATGCAACAGCAATATGTTTCTGAAACCAATCATT 1320
|
|
|
QY 1321 AAAGATCACACAGCTAGGCTTTCCTCGTAGCTTCCCTCTCTCTGCGATCTTTT 1380
|
|
|
Db 1321 AAAGATCACACAGCTAGGCTTTCCTCGTAGCTTCCCTCTCTCTGCGATCTTTT 1380
|
|
|
QY 1381 CGTCCACGACGATG 1394
|
|
|
Db 1381 CGTCCACGACGATG 1394
|
|
|

RESULT 2
US-08-880-499-2
Sequence 2, Application US/08880499
Patent No. 6037523
GENERAL INFORMATION:
APPLICANT: Albertson, Marc C.
APPLICANT: Fox, Tim W.
APPLICANT: Carl, Garnat W.
APPLICANT: Huffman, Gary A.
APPLICANT: Kendall, Timmy L.
TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION
NUMBER OF INVENTIONS: 2
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.
CITY: Johnston
STATE: Iowa
COUNTRY: USA
ZIP: 50131
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/880,499
FILING DATE: CONCURRENTLY HERewith
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Sweeney, Patricia A.
REGISTRATION NUMBER: 32,733
REFERENCE/DOCKET NUMBER: 0578
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 248-4800
TELEFAX: (515) 248-4844
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1394 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-880-499-2
Query Match 99.7%; Score 1389.2; DB 3; Length 1394;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1391; Conservative 0; Mismatches 3; Indels 0; Gaps 0;


```

; IMMEDIATE SOURCE:
; LIBRARY: maize endosperm cDNA library in
; CLONE: pMg110; pMg6Aa; pMg6C-2M
; US-08-968-542C-1
Query Match 4.8%; Score 66.4; DB 2; Length 6027;
Best Local Similarity 77.5%; Pred. No. 2.2e-06;
Matches 93; Conservative 0; Mismatches 26; Indels 1; Gaps 1;

QY 792 TTAAAGCTAGTTGGCAACCTGTTCTTCAAGATTGATTGTTTCAAAAAA 851
DB 5606 TCTACCGGCTAGTTGGGAACCCCATTT-TTCCAAGGATTTCCATTTCAGAAAA 5548
DB 5547 TTAGTTATTTCATTGGAAAAATGAAATCTTGGAAAAATAGATTCACTACTAG 5488

RESULT 4
US-09-554-467A-1/C
; Sequence 1, Application US/09554467A
; Patent No. 6639125
; GENERAL INFORMATION:
; APPLICANT: Myers, Alan M.
; APPLICANT: James, Martha G.
; TITLE OF INVENTION: dulla Coding for a No. 6639125el Starch Synthase and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: D6036PCT
; CURRENT APPLICATION NUMBER: US/09/554,467A
; CURRENT FILING DATE: 2000-05-12
; PRIOR APPLICATION NUMBER: PCT/US98/24225
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 08/062,102
; PRIOR FILING DATE: 1997-11-12
; NUMBER OF SEQ ID NOS: 37
; SEQ ID NO 1
; LENGTH: 6027
; TYPE: DNA
; ORGANISM: maize
; FEATURE:
; OTHER INFORMATION: cDNA sequence corresponding to the gene encoding the
; OTHER INFORMATION: starch synthase enzyme DUL.
; US-09-554-467A-1
Query Match 4.8%; Score 66.4; DB 4; Length 6027;
Best Local Similarity 77.5%; Pred. No. 2.2e-06;
Matches 93; Conservative 0; Mismatches 26; Indels 1; Gaps 1;

QY 792 TTAAAGCTAGTTGGCAACCTGTTCTTCAAGATTGATTGTTTCAAAAAA 851
DB 5606 TCTACCGGCTAGTTGGGAACCCCATTT-TTCCAAGGATTTCCATTTCAGAAAA 5548
DB 5547 TTAGTTATTTCATTGGAAAAATGAAATCTTGGAAAAATAGATTCACTACTAG 5488

RESULT 5
US-08-410-784A-3/C
; Sequence 3, Application US/08410784A
; Patent No. 5912413
; GENERAL INFORMATION:
; APPLICANT: MYERS, ALAN M.
; APPLICANT: JAMES, MARTHA G.
; TITLE OF INVENTION: ISOLATION OF SUL, A STARCH DEBRANCHING
; TITLE OF INVENTION: ENZYME, THE PRODUCT OF THE MAIZE GENE
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Weingarten, Schurjin, Gagnebin and Hayes LLP
; STREET: Ten Post Office Square
; CITY: Boston
```

```

; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/410,784A
; FILING DATE: 24-MAR-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Heine, Ph.D., Holliday C
; REGISTRATION NUMBER: 34,346
; REFERENCE/DOCKET NUMBER: ISU-002XX
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-2290
; TELEFAX: 617-451-0313
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2523 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; US-08-410-784A-3
Query Match 4.2%; Score 58.6; DB 2; Length 2523;
Best Local Similarity 71.4%; Pred. No. 0.00013;
Matches 105; Conservative 0; Mismatches 39; Indels 3; Gaps 2;

QY 774 AAAATCTTCTGATTTTAAAGCTAGTTGGCAACCTGTTCTTCAAGAAATT 833
DB 279 ACATATACATATACATATATATAGCTAGTTGCAATCCATTTTATCAAAAG--TTT 222
QY 834 TGATTTTCAAAAAAATTGTTATTTCTCTTATTAATAAATAGAAACACTGAAAA 893
DB 221 TACATTTTCAAAATTAATGTTATTTCTTGA-AAAATAGAAATTTCTCAGAAAA 163
DB 894 ATAGAGTTGCCAGACTAGCCCTAGAAAT 920
DB 162 ATAGAGTTTACAACCTAGCTTAAAT 136

RESULT 6
US-08-487-826B-13/C
; Sequence 13, Application US/08487826B
; Patent No. 5993827
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellem, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobb Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
```

```
ZIP: 92660
COMPUTER READABLE FORM.
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,825B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israelson, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121.001CP1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 19124 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHEetical: NO
ANTI-SENSE: NO
US-08-487-825B-13
```

```
Query Match 4.1%; Score 57; DB 2; Length 19124;
Best Local Similarity 50.2%; Pred. No. 0.00073;
Matches 141; Conservative 0; Mismatches 140; Indels 0; Gaps 0;
```

```
QY 614 TAAGCTATGACGCTTCGATTTTCTTTTCATCTGTGTAATTTGTTATGTTTATAT 673
DB 15677 TAATGTTTTTTTTTTCTCTCTTTGTTGTTTATTTTATATCATTTTATATAT 15618
QY 674 ACATTTCTCTCTTCAATAGAGTATTTCTTCGATTTTATTAAGCATATAAGT 733
DB 15617 AAAATTTTTTTTTTAAATTTTTTTTGATATCTTTTCATTTTATCTATCAAAATTTA 15558
QY 734 CATTTTATATTAAGACACGATGTCGTAGATTCGTCMAAAATCTTTCTGATTTTTT 793
DB 15557 TATTTTATATATATTTTATATTTTATTTTAAATAATTTCTCTTTTATTTTATTTT 15498
QY 794 TAAAGACTAGCTTGCGCAACCTGTTCTTTCAAAAGATTTTGATTTTCAAAAAAAT 853
DB 15497 TTTTATTTTAAATTAATTTTTTTTTTATATTTTCATTTTCTTTTATATTAATTT 15438
QY 854 AGTTATTTTCTCTTATTAATAATAGAAACACTAGAAAA 894
DB 15437 GTTTTATATTTCTTTTAAATTAATAATATATATATAA 15397
```

```
RESULT 7
US-09-313-294A-5397
Sequence 5397, Application US/09313294A
Patent No. 6476212
GENERAL INFORMATION:
APPLICANT: Lalsudi, Raghunath V.
APPLICANT: Ito, Laura Y.
APPLICANT: Sherman, Bradley K.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
FILE REFERENCE: PL-0017 US
CURRENT APPLICATION NUMBER: US/09/313,294A
NUMBER OF SEQ ID NOS: 1999-05-14
SOFTWARE: PERL Program
SEQ ID NO 5397
LENGTH: 279
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc_feature
```

```
OTHER INFORMATION: Incyte ID No. 6476212 700350078H1
NAME/KEY: unsure
LOCATION: 10, 12, 185-186, 204, 253, 274, 278
OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-5397
```

```
Query Match 4.0%; Score 55.8; DB 4; Length 279;
Best Local Similarity 69.0%; Pred. No. 0.00029;
Matches 89; Conservative 0; Mismatches 39; Indels 1; Gaps 1;
```

```
QY 789 TTTTAAAGAGTAGTTTGCAACCTGTTTCTTCAAGAAATTTGATTTTCAAAA 848
DB 3 TTCTTAANNCCTAGTTTGGAAACCCATTTTCCCAAGGTTTTCATTTTCCCAAGG 62
QY 849 AAATTAAGTTTATTTCTCTTAT-AAAATAGAAACACTTAGAAAAATAGAGTTGCCAGA 907
DB 63 AAGTTAGAACATTTTCCCTTGGGAAATAGAAATCTTTGGGAAATGAGTTCCCAA 122
QY 908 CTAGCCCTA 916
DB 123 CTAGCCCTA 131
```

```
RESULT 8
US-09-902-540-1357/C
Sequence 1357, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 1357
LENGTH: 612
TYPE: DNA
ORGANISM: Myxococcus xanthus
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(612)
OTHER INFORMATION: unsure at all n locations
US-09-902-540-1357
```

```
Query Match 4.0%; Score 55.6; DB 4; Length 612;
Best Local Similarity 51.4%; Pred. No. 0.00043;
Matches 127; Conservative 0; Mismatches 120; Indels 0; Gaps 0;
```

```
QY 626 TTCAATTTTCTTTTCAATCTGTATTTGTTATTTGTTTATATATATTTCTCT 685
DB 560 TTTTATATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 501
QY 686 CTACATAGAGTAGTTTCTTCGATTTTATTAATAAGCATTAATTTATATA 745
DB 500 TTTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATA 441
QY 746 AGAGCAGCATGTCGATGATCTCGTCAAAATCTTCTGATTTTAAAGACTAGTT 805
DB 440 ATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATATATTTATTT 381
QY 806 TGGCAACCTGTTTCTTCAAGAAATTTGATTTTCAAAAAATAGTTATTTTCT 865
DB 380 TTTTATTTTTTTTTTTTTTTTATTTATTTTATTTTATTTTATTTTATTTT 321
QY 866 CTTTATA 872
DB 320 TATTTTA 314
```


LENGTH: 396
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(396)
OTHER INFORMATION: n = A,T,C or G
US-09-640-173-53

Query Match 3.9%; Score 55; DB 4; Length 396;
Best Local Similarity 43.8%; Pred. No. 0.00052;
Matches 139; Conservative 0; Mismatches 178; Indels 0; Gaps 0;

QY 632 TTTTCTTTCTTCATCTGTTATTTGTTATTTGTTTAAATACATTTCTCTTACA 691
DB 11 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 70
QY 692 ATAGAGTGAATTTCTCCGATTTTAAATGACTATAAGTCAATTTTATATAGACA 751
DB 71 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 130
QY 752 CGCATGTCGATCTCTGTTCAAAATCTTCTGATTTTAAAGACTAGTTGGCAA 811
DB 131 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 190
QY 812 CCGTGTCTTCAAGAAATTTGATTTTTCAAAAAATAGTATTTCTCTTAT 871
DB 191 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 250
QY 872 AAAATGAAAAACCTTAGAAAAATAGAGTTGCCAGACTAGCCCTAGATGTTTCCCAAT 931
DB 251 AATTCANAAAAAGAAANAATTAANNANNCNNNNNNNNNNATNTNCTTNATA 310
QY 932 AAATTACATCACTGTG 948
DB 311 NTNNTNNNNNNANGG 327

RESULT 14
US-09-713-550-53
Sequence 53, Application US/09713550
Patent No. 6617109
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Stolc, John A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
THERAPY AND DIAGNOSIS OF OVARIAN CANCER
FILE REFERENCE: 210121.484C4
CURRENT APPLICATION NUMBER: US/09/713,550
CURRENT FILING DATE: 2000-11-14
NUMBER OF SEQ ID NOS: 205
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 53
LENGTH: 396
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(396)
OTHER INFORMATION: n = A,T,C or G
US-09-713-550-53

Query Match 3.9%; Score 55; DB 4; Length 396;
Best Local Similarity 43.8%; Pred. No. 0.00052;
Matches 139; Conservative 0; Mismatches 178; Indels 0; Gaps 0;

QY 632 TTTTCTTTCTTCATCTGTTATTTGTTATTTGTTTAAATACATTTCTCTTACA 691
DB 11 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 70
QY 692 ATAGAGTGAATTTCTCCGATTTTAAATGACTATAAGTCAATTTTATATAGACA 751
DB 71 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 130

QY 752 CGCATGTCGATCTCTGTTCAAAATCTTCTGATTTTAAAGACTAGTTGGCAA 811
DB 131 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 190
QY 812 CCGTGTCTTCAAGAAATTTGATTTTTCAAAAAATAGTATTTCTCTTAT 871
DB 191 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 250
QY 872 AAAATGAAAAACCTTAGAAAAATAGAGTTGCCAGACTAGCCCTAGATGTTTCCCAAT 931
DB 251 AATTCANAAAAAGAAANAATTAANNANNCNNNNNNNNNNATNTNCTTNATA 310
QY 932 AAATTACATCACTGTG 948
DB 311 NTNNTNNNNNNANGG 327

RESULT 15
US-09-825-294-53
Sequence 53, Application US/09825294
Patent No. 6710170
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Stolc, John A.
APPLICANT: Algate, Paul A.
APPLICANT: Fling, Steven P.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
THERAPY AND DIAGNOSIS OF OVARIAN CANCER
FILE REFERENCE: 210121.484C5
CURRENT APPLICATION NUMBER: US/09/825,294
CURRENT FILING DATE: 2001-04-03
NUMBER OF SEQ ID NOS: 215
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 53
LENGTH: 396
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(396)
OTHER INFORMATION: n = A,T,C or G
US-09-825-294-53

Query Match 3.9%; Score 55; DB 4; Length 396;
Best Local Similarity 43.8%; Pred. No. 0.00052;
Matches 139; Conservative 0; Mismatches 178; Indels 0; Gaps 0;

QY 632 TTTTCTTTCTTCATCTGTTATTTGTTATTTGTTTAAATACATTTCTCTTACA 691
DB 11 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 70
QY 692 ATAGAGTGAATTTCTCCGATTTTAAATGACTATAAGTCAATTTTATATAGACA 751
DB 71 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 130
QY 752 CGCATGTCGATCTCTGTTCAAAATCTTCTGATTTTAAAGACTAGTTGGCAA 811
DB 131 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 190
QY 812 CCGTGTCTTCAAGAAATTTGATTTTTCAAAAAATAGTATTTCTCTTAT 871
DB 191 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 250
QY 872 AAAATGAAAAACCTTAGAAAAATAGAGTTGCCAGACTAGCCCTAGATGTTTCCCAAT 931
DB 251 AATTCANAAAAAGAAANAATTAANNANNCNNNNNNNNNNATNTNCTTNATA 310
QY 932 AAATTACATCACTGTG 948
DB 311 NTNNTNNNNNNANGG 327

Search completed: September 15, 2005, 08:24:54
Job time : 214.859 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 07:10:11 ; Search time 907.859 Seconds
(without alignments)
10230.248 Million cell updates/sec

Title: US-10-713-381-1

Perfect score: 1394

Sequence: 1 ccatggtcgtctctatgaaaaa.....cttttcgtccacaccacatg 1394

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 7389322 seqs, 333185599 residues

Total number of hits satisfying chosen parameters: 14778644

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/prodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/prodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/prodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/prodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/prodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/prodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/prodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/prodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/prodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/prodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/prodata/2/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/prodata/2/pubpna/US10G_PUBCOMB.seq:*
20: /cgn2_6/prodata/2/pubpna/US10H_PUBCOMB.seq:*
21: /cgn2_6/prodata/2/pubpna/US10I_PUBCOMB.seq:*
22: /cgn2_6/prodata/2/pubpna/US10_NEW_PUB.seq:*
23: /cgn2_6/prodata/2/pubpna/US11A_PUBCOMB.seq:*
24: /cgn2_6/prodata/2/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1394	100.0	1394	20	US-10-713-381-1
2	1389.2	99.7	1394	20	US-10-713-381-2
3	238.8	17.1	225	20	US-10-713-381-9
4	146	10.5	158	20	US-10-713-381-3
5	80.8	5.8	1261	20	US-10-425-115-134230
6	78.8	5.7	1663	20	US-10-739-930-1857
7	74	5.3	320	18	US-10-425-114-23340

8	74	5.3	624	18	US-10-425-114-16264	Sequence 16264, A
9	73	5.2	1326	20	US-10-425-115-141826	Sequence 141826, A
10	72.8	5.2	2445	18	US-10-425-114-32493	Sequence 32493, A
11	72.8	5.2	2729	20	US-10-425-115-83293	Sequence 83293, A
12	72.4	5.2	1203	20	US-10-425-115-51470	Sequence 51470, A
13	71.8	5.2	928	18	US-10-425-114-17816	Sequence 17816, A
14	71.8	5.2	1123	20	US-10-425-115-142853	Sequence 142853, A
15	71.8	5.2	2537	18	US-10-425-114-11957	Sequence 31957, A
16	71.8	5.2	2537	20	US-10-425-115-52216	Sequence 52216, A
17	71.8	5.2	3607	18	US-10-425-114-31061	Sequence 31061, A
18	71.8	5.2	3691	20	US-10-425-115-52219	Sequence 52219, A
19	70.2	5.0	960	14	US-10-198-846-6381	Sequence 6381, Ap
20	69.8	5.0	13400	21	US-10-343-477A-63	Sequence 63, Ap1
21	69.6	5.0	724	20	US-10-425-115-87756	Sequence 87756, A
22	69.2	5.0	527	18	US-10-425-114-4041	Sequence 4041, Ap
23	68.8	4.9	610	20	US-10-425-115-47165	Sequence 47165, A
24	68.4	4.9	1215	18	US-10-425-114-24656	Sequence 24656, A
25	68.4	4.9	2863	20	US-10-425-115-75310	Sequence 75310, A
26	67.8	4.9	512	20	US-10-425-115-49781	Sequence 49781, A
27	67.8	4.9	928	18	US-10-425-114-17816	Sequence 17816, A
28	67.8	4.9	1123	20	US-10-425-115-142853	Sequence 142853, A
29	67.4	4.8	1260	18	US-10-425-114-30881	Sequence 30881, A
30	67.4	4.8	1411	20	US-10-425-115-149304	Sequence 149304, A
31	67.2	4.8	2232	18	US-10-425-114-16607	Sequence 16607, A
32	67.2	4.8	2249	18	US-10-425-114-30264	Sequence 20264, A
33	66.4	4.8	6027	18	US-10-634-262-1	Sequence 1, Appli
34	66.4	4.8	6027	19	US-10-109-048-1145	Sequence 1145, Ap
35	66.4	4.8	6051	20	US-10-425-115-41589	Sequence 41589, A
36	66.2	4.7	439	9	US-09-960-352-1009	Sequence 1009, Ap
37	66	4.7	1376	18	US-10-425-114-2073	Sequence 2073, Ap
38	65.6	4.7	721	20	US-10-425-115-177935	Sequence 177935, Ap
39	65.6	4.7	2274	18	US-10-425-114-777	Sequence 777, App
40	65.6	4.7	2274	18	US-10-425-114-830	Sequence 830, App
41	65.6	4.7	2832	20	US-10-425-115-162957	Sequence 162957, A
42	65.6	4.7	8056	20	US-10-475-126-186	Sequence 386, App
43	65	4.7	1364	20	US-10-425-115-150820	Sequence 150820, A
44	64.8	4.6	833	20	US-10-425-115-30824	Sequence 30824, A
45	64.8	4.6	1838	20	US-10-425-115-21012	Sequence 21012, A

ALIGNMENTS

RESULT 1
US-10-713-381-1
; Sequence 1, Application US/10713381
; Publication No. US2004022131A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNMAN, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-1

Query Match 100.0%; Score 1394; DB 20; Length 1394;
Best Local Similarity 100.0%; Pred. No. 1,16-294;
Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCATGTCGTCTCTATGAAAAGATAGTACATGTGTCTATTCGTTTCTTAGGCTCC 60

Db 1 CCAATGCTCTATGAAAAAGATGAGTCAATGTCCTATATCCGTTTCTTAAGGTC 60
Qy 61 CTTCTTCGCTTATTAAGTCAATCGGGTTTCAAAAACTTCAAGGTCATGAT 120
Db 61 CTTCTTCGCTTATTAAGTCAATCGGGTTTCAAAAACTTCAAGGTCATGAT 120
Qy 121 CTCATGTTCACTCTCCACCTGCGTGGCACTTCTTGGATGCGGTGGTCCAT 180
Db 121 CTCATGTTCACTCTCCACCTGCGTGGCACTTCTTGGATGCGGTGGTCCAT 180
Qy 181 CTGACCGAGGCCATCGACACCTTTCGGGACCCATCAAGGCGCTTTCGATGGCCA 240
Db 181 CTGACCGAGGCCATCGACACCTTTCGGGACCCATCAAGGCGCTTTCGATGGCCA 240
Qy 241 CGAGAGTATCGGTCGTGTGATCCAGGGGATATATGTCCCGACATGTCACCTATA 300
Db 241 CGAGAGTATCGGTCGTGTGATCCAGGGGATATATGTCCCGACATGTCACCTATA 300
Qy 301 TTATATTTCTTATGATATATTTTGAAGAAAAATACAACTTATCTTTGTGTA 360
Db 301 TTATATTTCTTATGATATATTTTGAAGAAAAATACAACTTATCTTTGTGTA 360
Qy 361 GGGCTTCAGATGATTTTGGCTTGAAGCCAGAAATGCGAGGACGACCATGTCATG 420
Db 361 GGGCTTCAGATGATTTTGGCTTGAAGCCAGAAATGCGAGGACGACCATGTCATG 420
Qy 421 TCCATTTTGGCATCCCGCAAGATTTAAATAAACAAGTAACTTACCT 480
Db 421 TCCATTTTGGCATCCCGCAAGATTTAAATAAACAAGTAACTTACCT 480
Qy 481 CGAAGCTATCATGATATGTTTAAAGAAATCATATTTAAACGACATCCTTTAAAA 540
Db 481 CGAAGCTATCATGATATGTTTAAAGAAATCATATTTAAACGACATCCTTTAAAA 540
Qy 541 CAAGCATATTTGAAAGAGCAAAATATGTTACAGTTTCAAAATCTTAAGCGACA 600
Db 541 CAAGCATATTTGAAAGAGCAAAATATGTTACAGTTTCAAAATCTTAAGCGACA 600
Qy 601 TTTATCGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Db 601 TTTATCGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Qy 661 ATTGTTTTTATATATATTTCTCTTCAATAGATGATTTCTTCCGATTTTAA 720
Db 661 ATTGTTTTTATATATATTTCTCTTCAATAGATGATTTCTTCCGATTTTAA 720
Qy 721 ATGACTATTAAGTCAATTTTATATTAAGAGCAGCATGTCGATGATTTCTGTC 780
Db 721 ATGACTATTAAGTCAATTTTATATTAAGAGCAGCATGTCGATGATTTCTGTC 780
Qy 781 TTTCTGATTTTAAAGCTAGTTGGCAACCTGTTCTTCAAGAAATTTGATTT 840
Db 781 TTTCTGATTTTAAAGCTAGTTGGCAACCTGTTCTTCAAGAAATTTGATTT 840
Qy 841 TTCAAAAAAATAGTTTATTTCTCTTATTAATAAGAAATCACTTAAAGAAAT 900
Db 841 TTCAAAAAAATAGTTTATTTCTCTTATTAATAAGAAATCACTTAAAGAAAT 900
Qy 901 TGCAGACTAGCCCTGAGATGTTTCCCAATTAATCACTGTAATTAATTTG 960
Db 901 TGCAGACTAGCCCTGAGATGTTTCCCAATTAATCACTGTAATTAATTTG 960
Qy 961 GCGAGCCCAATTAATTTAAACGAAATGAAATCGAGCGCAACCAATCTGACCT 1020
Db 961 GCGAGCCCAATTAATTTAAACGAAATGAAATCGAGCGCAACCAATCTGACCT 1020
Qy 1021 TTCTCTAGATTAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Db 1021 TTCTCTAGATTAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Qy 1081 TGAGATGTCGGTTTGGCAAGATGACCGATATGATGATGATGATGATGATGAT 1140
Db 1081 TGAGATGTCGGTTTGGCAAGATGACCGATATGATGATGATGATGATGATGAT 1140

Db 1081 TGAGATGTCGGTTTGGCAAGATGACCGATATGATGATGATGATGATGAT 1140
Qy 1141 GGTTCGAGCTCTGTCATCTCATGATGATGATGATGATGATGATGATGAT 1200
Db 1141 GGTTCGAGCTCTGTCATCTCATGATGATGATGATGATGATGATGATGAT 1200
Qy 1201 TTGTTCAATGTCGAGCTTCTCTATTTCTGAACAGAGATTAATCTTCCAA 1260
Db 1201 TTGTTCAATGTCGAGCTTCTCTATTTCTGAACAGAGATTAATCTTCCAA 1260
Qy 1261 CCATCTTACTCATGCACTTCCATGCAACGCAATATGTTTCTGAACCAATCA 1320
Db 1261 CCATCTTACTCATGCACTTCCATGCAACGCAATATGTTTCTGAACCAATCA 1320
Qy 1321 AAGATGCAACAGCTAGGTTCTCCGCTAGCTTCCCTCTCTCTGCGATCT 1380
Db 1321 AAGATGCAACAGCTAGGTTCTCCGCTAGCTTCCCTCTCTCTGCGATCT 1380
Qy 1381 CGTCCACGACATG 1394
Db 1381 CGTCCACGACATG 1394

RESULT 2
US-10-713-381-2
; Sequence 2, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: USING SAME
; CURRENT FILING DATE: 2003-11-14
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-2

Query Match 99.7%; Score 1389.2; DB 20; Length 1394;
Best Local Similarity 99.8%; Pred. No. 1.2e-293;
Matches 1391; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CCATGCTCTCTATGAAAAAGATGAGTCAATGTCCTATATCCGTTTCTTAAGGTC 60
Db 1 CCATGCTCTCTATGAAAAAGATGAGTCAATGTCCTATATCCGTTTCTTAAGGTC 60
Qy 61 CTTCTTCGCTTATTAAGTCAATCGGGTTTCAAAAACTTCAAGGTCATGAT 120
Db 61 CTTCTTCGCTTATTAAGTCAATCGGGTTTCAAAAACTTCAAGGTCATGAT 120
Qy 121 CTCATGTTCACTCTCCACCTGCGTGGCACTTCTTGGATGCGGTGGTCCAT 180
Db 121 CTCATGTTCACTCTCCACCTGCGTGGCACTTCTTGGATGCGGTGGTCCAT 180
Qy 181 CTGACCGAGGCCATCGACACCTTTCGGGACCCATCAAGGCGCTTTCGATGGCCA 240
Db 181 CTGACCGAGGCCATCGACACCTTTCGGGACCCATCAAGGCGCTTTCGATGGCCA 240
Qy 241 CGAGAGTATCGGTCGTGTGATCCAGGGGATATATGTCCCGACATGTCACCTATA 300
Db 241 CGAGAGTATCGGTCGTGTGATCCAGGGGATATATGTCCCGACATGTCACCTATA 300
Qy 301 TTATATTTCTTATGATATATTTTGAAGAAAAATACAACTTATCTTTGTGTA 360
Db 301 TTATATTTCTTATGATATATTTTGAAGAAAAATACAACTTATCTTTGTGTA 360

Db 301 TTATTTCTTTAGATTTATTTATTTTGGAAAAATTAACAACTTATCTTTGTGTA 360
361 GGGCTTCAGCATAGATTTTCGCTTAGGGCCAGAAATCGAGGACGACCATGTCTAGT 420
Db 361 GGGCTTCAGCATAGATTTTCGCTTAGGGCCAGAAATCGAGGACGACCATGTCTAGT 420
Qy 421 TCCACTTTGGCACTACCCAGAACAGATTTAAAAATAACCAAGTAATCTATCCACT 480
Db 421 TCCACTTTGGCACTACCCAGAACAGATTTAAAAATAACCAAGTAATCTATCCACT 480
Qy 481 CGAAGCTATCATGTATGTTTAAAGAAACATCTATTAACCAAGTCCCTTAAAAAA 540
Db 481 CGAAGCTATCATGTATGTTTAAAGAAACATCTATTAACCAAGTCCCTTAAAAAA 540
Qy 541 CAAGCATATTTTGAAGAGACAAATTTATGTTACAGTTTACAAATCTTAAGAGCGACAA 600
Db 541 CAAGCATATTTTGAAGAGACAAATTTATGTTACAGTTTACAAATCTTAAGAGCGACAA 600
Qy 601 TTTATATGAAAGGTAGCTATGACGTTCAAGATTTTCTTTTCAATTTCTGTAATTTGTT 660
Db 601 TTTATATGAAAGGTAGCTATGACGTTCAAGATTTTCTTTTCAATTTCTGTAATTTGTT 660
Qy 661 ATTGTTTTATATACATTTTCTCTCTTACAAATAGAGATTTTCTCGATTTTAA 720
Db 661 ATTGTTTTATATACATTTTCTCTCTTACAAATAGAGATTTTCTCGATTTTAA 720
Qy 721 ATGACTATAAAGTCATTTTATATATAGAGACGCGATGCTGTAATCTCGTTCAAAATC 780
Db 721 ATGACTATAAAGTCATTTTATATATAGAGACGCGATGCTGTAATCTCGTTCAAAATC 780
Qy 781 TTTCTGATTTTTTAAAGCTAGTTTGGCAACCTGTTCTTTCAAGAAATTTGATTTT 840
Db 781 TTTCTGATTTTTTAAAGCTAGTTTGGCAACCTGTTCTTTCAAGAAATTTGATTTT 840
Qy 841 TTTCAAAAAAATAGTTTATTTCTTTTAAATAAGAAACATTAAGAAATTTGAGT 900
Db 841 TTTCAAAAAAATAGTTTATTTCTTTTAAATAAGAAACATTAAGAAATTTGAGT 900
Qy 901 TGGCAGACTAGCCCTAGAAATGTTTTCCAAATAATTAACAATCACTGTAATAATTTT 960
Db 901 TGGCAGACTAGCCCTAGAAATGTTTTCCAAATAATTAACAATCACTGTAATAATTTT 960
Qy 961 GCCAGCCCCATTAATTTTAAACCGAACTGAATGAGCGAAACCAATCTGAGCTAT 1020
Db 961 GCCAGCCCCATTAATTTTAAACCGAACTGAATGAGCGAAACCAATCTGAGCTAT 1020
Qy 1021 TTTCTAGATTTAGTAAAGAGAGAGAGAGAAATCAGTTTAACTCATTTGCC 1080
Db 1021 TTTCTAGATTTAGTAAAGAGAGAGAGAGAAATCAGTTTAACTCATTTGCC 1080
Qy 1081 TGAGATGTGCGTTTGGCAAGATAGCCAGTAACTCATAGTGCTTACGTC 1140
Db 1081 TGAGATGTGCGTTTGGCAAGATAGCCAGTAACTCATAGTGCTTACGTC 1140
Qy 1141 GGTTCGGCACTCTCGTGCATCTCACATGCACTACTAGTGTGTTCAACGTTGTC 1200
Db 1141 GGTTCGGCACTCTCGTGCATCTCACATGCACTACTAGTGTGTTCAACGTTGTC 1200
Qy 1201 TTGTTTCATGCTGCAAGCTTGGCTATTTGAAACCAAGAGATACCACTCCCAACAT 1260
Db 1201 TTGTTTCATGCTGCAAGCTTGGCTATTTGAAACCAAGAGATACCACTCCCAACAT 1260
Qy 1261 CCATCTTACTCATGCACTTTCATGCAAGACGACATATGTTTCCGAAACAGATCAT 1320
Db 1261 CCATCTTACTCATGCACTTTCATGCAAGACGACATATGTTTCCGAAACAGATCAT 1320
Qy 1321 AAAGATCAACAAGCTAGGCTTTCGCGTAGCTTCCCTCTCTCTGCGCATCTTTT 1380
Db 1321 AAAGATCAACAAGCTAGGCTTTCGCGTAGCTTCCCTCTCTCTGCGCATCTTTT 1380
Qy 1381 CGTCACACCATG 1394
Db 1381 CGTCACACCATG 1394

Db 1381 CGTCACACCATG 1394
RESULT 3
US-10-713-381-9
; Sequence 9, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFEMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 255
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-9
Query Match 17.1%; Score 238.8; DB 20; Length 255;
Best Local Similarity 97.2%; Pred. No. 3.2e-42;
Matches 243; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy 1145 CGGCACTCTGCTGATCTCAATGCACTACTACATGCTTGTCAACGTTGCTGT 1204
Db 5 CCGCGATCCCGTGATCTCAATGCACTACTACATGCTTGTCAACGTTGCTGT 64
Qy 1205 TCCATGTCGAAGCTTGGCTTATTTGAAACCAAGAGATACCTACTTCCAAACATCAT 1264
Db 65 TCCATGTCGAAGCTTGGCTTATTTGAAACCAAGAGATACCTACTTCCAAACATCAT 124
Qy 1265 CTTACTCATGCACTTCAATGCAAGACGACATATGTTTCTGAAACCAATCATTAAG 1324
Db 125 CTTACTCATGCACTTCAATGCAAGACGACATATGTTTCTGAAACATCATTAAG 184
Qy 1325 ATCAACAAGCTAGCGTTTCCGCTAGCTTCCCTCTCTCTGCGCATCTTTTCTGTC 1384
Db 185 ATCAACAAGCTAGCGTTTCCGCTAGCTTCCCTCTCTCTGCGCATCTTTTCTGTC 244
Qy 1385 CACCAACATG 1394
Db 245 CACCAACATG 254
RESULT 4
US-10-713-381-3
; Sequence 3, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFEMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3

```
/ LENGTH: 158
/ TYPE: DNA
/ ORGANISM: Zea mays
US-10-713-381-3

Query Match      10.5%; Score 146; DB 20; Length 158;
Best Local Similarity 99.4%; Pred. No. 5.3e-22;
Matches 157; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1155 CGGTGATCTCATGCGCATGCAATGCTGTTGCAACCGTTGTC-TTGTTCATGCTC 1213
DB 1 CGGTGATCTCATGCGCATGCAATGCTGTTGCAACCGTTGTTGTTCCATGCTC 60

QY 1214 CAAGCCTTGCTTATTTCTGAACCAAGAGATACCTACCCCAACATTCATCTTACTCAT 1273
DB 61 CAAGCCTTGCTTATTTCTGAACCAAGAGATACCTACCCCAACATTCATCTTACTCAT 120

QY 1274 GCAACTTCATGCAACGCAACGCAATATGTTTCTGAAAC 1311
DB 121 GCAACTTCATGCAACGCAACGCAATATGTTTCTGAAAC 158

RESULT 5
US-10-425-115-134230
/ Sequence 134230, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21 (53222)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 134230
/ LENGTH: 1261
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MRT4577_53903C.1
US-10-425-115-134230

Query Match      5.8%; Score 80.8; DB 20; Length 1261;
Best Local Similarity 78.2%; Pred. No. 2.7e-07;
Matches 97; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 792 TTTAAGAGCTAGTTTGCAACCGCTGTTCTTCAAGAAATTTGATTTTCAAAAAA 851
DB 999 TCTTAAGGCTAGTTTGGAACCTTATTTTAAAGATTTTCTTTTTTAAAGATAA 1058

QY 852 TTAGTTTATTTCTCTTAAATAATAGAAAACACTTAGAAAAATAGAGTTGCCAGACTAG 911
DB 1059 TTAGTTTATTTCTCTTGAATAATAGAAATCTCTGGAATAATGAGAGTTGCTAACTAG 1118

QY 912 CCCT 915
DB 1119 CCCT 1122

RESULT 6
US-10-739-930-1857
/ Sequence 1857, Application US/10739930
/ Publication No. US20040216190A1
/ GENERAL INFORMATION:
/ APPLICANT: Kovalic, David K.
/ TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
/ FILE REFERENCE: 38-21 (53377)B
/ CURRENT APPLICATION NUMBER: US/10/739,930
/ CURRENT FILING DATE: 2003-12-18
```

```
/ NUMBER OF SEQ ID NOS: 11088
/ SEQ ID NO 1857
/ LENGTH: 1663
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: ZEAMA-23APR03-CLUSTER1429_8
US-10-739-930-1857

Query Match      5.7%; Score 78.8; DB 20; Length 1663;
Best Local Similarity 97.6%; Pred. No. 8.3e-07;
Matches 80; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1313 AATCCATTAAAGATCAACACAGCTAGCGTTCTCCGCTTCCCTCTCTCTGCGG 1372
DB 1 AATCCATTAAAGATCAACACAGCTAGCGTTCTCCGCTTCCCTCTCTCTGCGG 60

QY 1373 ATCTTTTGTCTCACCAACCATG 1394
DB 61 ATCTTTTGTCTCACCAACCATG 82

RESULT 7
US-10-425-114-23340
/ Sequence 23340, Application US/10425114
/ Publication No. US2004003488A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jingdong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E
/ APPLICANT: Tabaska, Jack E
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21 (53313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 23340
/ LENGTH: 320
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: LIB595-039-G8_FLI
US-10-425-114-23340

Query Match      5.3%; Score 74; DB 18; Length 320;
Best Local Similarity 73.1%; Pred. No. 4.2e-06;
Matches 95; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 791 TTTAAGAGCTAGTTTGCAACCGCTGTTCTTCAAGAAATTTGATTTTCAAAAAA 850
DB 62 TTTTAAGGCTAGTTTGGAATCTCATTTTCAAGAGATTTTATTTTCTTAAGAA 121

QY 851 ATTAGTTTATTTCTCTTAAATAATAGAAAACACTTAGAAAAATAGAGTTCCAGACTA 910
DB 122 ATTAGTTTATTTCTCTTGAATAATAGAAATCCCTTGGAAATTAGAGTTCTAAACA 181

QY 911 GCCCTAGAAAT 920
DB 182 GCCCTAAAT 191

RESULT 8
US-10-425-114-16264
/ Sequence 16264, Application US/10425114
/ Publication No. US2004003488A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jingdong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E
```

```
/ APPLICANT: Tabaska, Jack E
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(5313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 16264
/ LENGTH: 624
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: LIB3062-023-G10_FLI
US-10-425-114-16264
```

```
Query Match      5.3%; Score 74; DB 18; Length 624;
Best Local Similarity 73.1%; Pred. No. 5.8e-06;
Matches 95; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
```

```
QY 791 TTTTAAGAGCTAGTTGGCAACCCGTTCTTCAAGAAATTTGATTTTCAAAAAA 850
    |||||
DB 374 TTTTGGGCTAGTTGGAATCTATTTTCAAGAGATTTTCTTAAGAA 433
    |||||
QY 851 ATTAGTTATTTCTCTTTAAATAGAAACCTTAGAAATAGAGTTGCCAGACTA 910
    |||||
DB 434 ATTAGTTATTTCTCTTTGAGAAATAGAAATCCCTTGGAAATAGAGTTCTAAACGA 493
    |||||
QY 911 GCCCTAGAAAT 920
    |||||
DB 494 GCCCTTAATT 503
    |||||
```

```
RESULT 9
US-10-425-115-141826
/ Sequence 141826, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(53222)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 141826
/ LENGTH: 1326
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MRT4577_60829C.1
US-10-425-115-141826
```

```
Query Match      5.2%; Score 73; DB 20; Length 1326;
Best Local Similarity 69.1%; Pred. No. 1.4e-05;
Matches 114; Conservative 0; Mismatches 50; Indels 1; Gaps 1;
```

```
QY 756 TGTCGAGATTCCTCGTTCAAAAAATCTTGTGATTTTAAAGCTAGTTGGCAACCT 815
    |||||
DB 1090 TGAAGTGGGTGGTGGCTGATTCACCTTTAACTGTAAGACTAATTTGGAAACCA 1149
    |||||
QY 816 GTTCTTTCAAGAAATTTGATTTTCAAAAAATAGTTATTTCTTTATPAAA 875
    |||||
DB 1150 ATTT-TTCACTGATTTTCAATTTCTTAAGAAATAGTTATTTCCCTTGAGAAAA 1208
    |||||
QY 876 TAGAAACACTTAGAAATAGAGTTGCCAGACTAGCCCTAGAAAT 920
    |||||
DB 1209 TAGGAATCCCTTAGAAAAAATAGTTTCCAACTAGCCCTTAAT 1253
    |||||
```

```
RESULT 10
US-10-425-114-32493
/ Sequence 32493, Application US/10425114
/ Publication No. US20040034888A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jingdong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E
/ APPLICANT: Tabaska, Jack E
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(5313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 32493
/ LENGTH: 2445
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: UC-ZMFLB73402B09_FLI
US-10-425-114-32493
```

```
Query Match      5.2%; Score 72.8; DB 18; Length 2445;
Best Local Similarity 73.6%; Pred. No. 2.1e-05;
Matches 106; Conservative 0; Mismatches 37; Indels 1; Gaps 1;
```

```
QY 792 TTTAAGAGCTAGTTGGCAACCCGTTCTT-TCAAGAAATTTGATTTTCAAAAAA 850
    |||||
DB 1623 TCTTGGGCTAGTTGAGAACCCCTTTTTCCTCAAGATTTTCAAGAAAA 1682
    |||||
QY 851 ATTAGTTATTTCTCTTTAAATAGAAACACTTAGAAATAGAGTTGCCAGACTA 910
    |||||
DB 1683 ATTAGTTATTTCTCTTTGAGAAATAGAAATCCCTTAABAAAAATAGTTTCAAACTA 1742
    |||||
QY 911 GCCCTAGAAATGTTTCCCAATPAA 934
    |||||
DB 1743 GCCCTTAATGTTTTCATGAA 1766
    |||||
```

```
RESULT 11
US-10-425-115-83293/C
/ Sequence 83293, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(53222)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 83293
/ LENGTH: 2729
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (1)..(2729)
/ OTHER INFORMATION: unsure at all n locations
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MRT4577_175978C.1
US-10-425-115-83293
```

```
Query Match      5.2%; Score 72.8; DB 20; Length 2729;
Best Local Similarity 60.7%; Pred. No. 2.2e-05;
Matches 136; Conservative 0; Mismatches 87; Indels 1; Gaps 1;
```

QY	792	TTTAGACCTAGTTGGCAACCCGTTCTTCCAAAGATTGTATTTTTCAAAAAAA	851
Db	2553	TTTAAGGGCTAGTTTAGGA--CCTATTTTTTCTAGAGTTTCTATTTTTCACGAAAA	2465
QY	852	TTAGTTTATTTTCTCTTATATAAATAGAAAAACATTAGAAAAATAGAGTGCACAGTAG	911
Db	2464	ATATTTCATTTCCCTTGGAAAAATAGAAATACCTTGGAAAAATATGAGTTTCCAAACTG	2405
QY	912	CCCTAGAAATGTTTCCCAATAATTACAATCATCTGTATATAATTATTTGGCCAGCCCAT	971
Db	2404	CCTTCAAAATTTCTCTCAAGTATTTTCTTCATAGAAAGTATGTATCAATGCGCACT	2345
QY	972	AAATTAATTTAAACGAAACTGAAATCGAGCGAAACCAATCTGA	1015
Db	2344	TGAGATCATGAGCTAAATATGTAAAGGATAAATCAATATGTTA	2301

```

RESULT 12
US-10-425-115-51470/C
; Sequence 51470, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 51470
; LENGTH: 1203
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1203)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_146936C.1
; US-10-425-115-51470

```

```

Query Match          5.2%; Score 72.4; DB 20; Length 1203;
Best Local Similarity 76.1%; Pred. No. 1.8e-05;
Matches 102; Conservative 0; Mismatches 31; Indels 1; Gaps 1;

Qy      787  ATTTTAAAGAGCTAGTTTGCAACCCGTGTTCTTTCAAAGATTTATATTTTTCANA 846
Db      1189  ACTTATTTAAGGCTTAGTGGGGA-CATATTTTTCACAAAAGATTTCTATTTTCTTAA 1131

Qy      847  AAAAAATGATTATTTTCTCTTTATTAATAATAGAAAAACATTAGAAAAATAGAGTCCNG 906
Db      1130  GAAATTTGTTTATTTTCTCTTGAAAAATAAATATCCGTGAAAAAATAGAGTTCCAA 1071

Qy      907  ACTAGCCCTAGAT 920
Db      1070  ACTAGCTCTAAGAT 1057

RESULT 13
US-10-425-114-17816/c
; Sequence 17816, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

```

```

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 17816
;
; LENGTH: 928
; TYPE: DNA
; ORGANISM: Zea mays
;
; FEATURE:
; OTHER INFORMATION: clone ID: LIB3075-032-A3_FLI
; US-10-425-114-17816
;
Query Match: 5.2%; Score 71.8; DB 18; Length 928;
Best Local Similarity 78.0%; Pred. No. 2,1e-05;
Matches 99; Conservative 0; Mismatches 27; Indels 1; Gaps 1;

```

[illegible]

```

RESULT 14
US-10-425-115-142853/C
; Sequence 142853, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 142853
; LENGTH: 1123
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_61761C.1
; US-10-425-115-142853

```

	Query Match	Similarity	5.2%	Score 71.8;	DB 20;	Length 1123;
	Best Local	Similarity	78.0%;	Fred. No. 2.3e-05;		
	Matches	Conservative	0;	Mismatches 27;	Indels 1;	Gaps 1
Qy	790	TTTTAAAGCTAGTTGGCAACCCGCTGTTCTTCAAGAAATTTGATTTTTCAAAAA	849			
Db	821	TTTTTAAGCTGTGTTGGCAA-CTCAATTTTCTAAGGATTTCTATTTTACTAAGAA	763			
Qy	850	AATTAGTTATTTTCTCTTTAATAAATAGAAAAACAATTAGAAAAATAGAGTCCAGACT	909			
Db	762	AATTAGTTAATTTTCTCTTAGAAAAATAAATATGCTTGGAAAAATAGAGTACCAAATT	703			
Qy	910	AGCCCTA 916				
Db	702	ACCCTTA 696				

RESULT 15
US-10-425-114-31957
; Sequence 31957, Application US/10425114


```
/ Publication No. US20040034888A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jingdong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E
/ APPLICANT: Tabaska, Jack E
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(5313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ. ID NOS: 73128
/ SEQ. ID NO 31957
/ LENGTH: 2537
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: UC-ZMFLB73238H12_FLI
US-10-425-114-31957

Query Match      5.2%; Score 71.8; DB 18; Length 2537;
Best Local Similarity 73.4%; Pred. No. 3.5e-05;
Matches 105; Conservative 0; Mismatches 37; Indels 1; Gaps 1;

QY      792 TTTAAGAGCTAGTTGGCAACCTGTTCTTTCAGAGATTTTGATTTTCAAAAAA 851
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      1799 TCTTAGGGCTAGTTGGAACCT-TTTTCCCAAAAGATTTCATTTTCAAGAAAA 1857

QY      852 TTAGTTATTTCTCTTATAAATGAAACCTTAGAAATAGAGTGCAGACTAG 911
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      1858 TTAGTTCAATTTTCTTCTTGAGAAATAGGAATCCCTTAAAAAATAGTTCAAACTAG 1917

QY      912 CCTAGATGTTTCCCATPAA 934
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      1918 CCCTTAATGTTTTCATGAA 1940
```

Search completed: September 15, 2005, 20:45:31
Job time : 909.859 secs

This Page Blank (unfiled)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using bw model

Run on: September 15, 2005, 16:07:32 ; Search time 4347.69 Seconds

(without alignments)
12204.542 Million cell updates/sec

Title: US-10-713-381-1

Perfect score: 1394
Sequence: 1 cccttggtgcctcctcgaataa.....cttttcgctccaccaccatg 1394

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hnc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_g881:*
9: gb_g882:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	952	68.3	963	9	CC656933 OGMDO20TM
2	760.4	54.5	915	9	CG224225 OGIAG08TV
3	501.2	36.0	687	9	CC656939 OGMDO20TV
4	96	6.9	715	9	CG252571 OGB805TC
5	92	6.6	967	9	CL235046 ZMMBB057
6	88	6.3	814	9	CG048704 PU1019TB
7	87.8	6.3	754	9	CG414922 ZMMBB029
8	86.8	6.2	950	8	CC439901 PUHVI5TB
9	86.6	6.2	652	8	CC384247 PUHOC67TB
10	86.6	6.2	797	8	CC400575 PUHLU61TD
11	86.6	6.2	820	8	CC400574 PUHLU61TB
12	85.4	6.1	471	9	CG103452 PUHBI9TB
13	85.2	6.1	765	9	CG082135 PUFOX12TD
14	85.2	6.1	781	9	CG630219 OGU053TV
15	85.2	6.1	815	9	CG349565 OGU0F183TV
16	85.2	6.1	834	9	CC630210 OGU053TV
17	85	6.1	789	8	CC433618 PUHPI7TD
18	84.6	6.1	1092	8	CNS020K7
19	84	6.0	793	8	BZ816381
20	83.8	6.0	1078	9	CL997678
21	83.6	6.0	530	9	CG201774
22	83.6	6.0	722	9	CG333914
23	83.6	6.0	722	9	CG333929
24	83.6	6.0	781	9	CG034985

25	83.6	6.0	861	8	BZ797976
26	83.6	6.0	865	8	CC430754
27	81.8	5.9	861	9	CG102092
28	81.8	5.9	947	8	CC435780
29	81.8	5.9	981	8	BZ784278
30	81.8	5.9	1017	9	CC620594
31	81.6	5.9	649	9	CC613918
32	81.6	5.9	733	8	BZ778636
33	81.6	5.9	925	9	CG071791
34	81.6	5.9	960	8	BZ676889
35	81.6	5.9	994	8	CC003943
36	81.6	5.9	1016	9	CL996481
37	81.6	5.9	1022	9	CL984151
38	81.4	5.8	999	8	CC385762
39	81.2	5.8	756	9	CG221693
40	81	5.8	1101	9	CNS00172
41	80.6	5.8	699	8	BZ96930
42	80.6	5.8	832	9	CC676387
43	80.6	5.8	839	9	CG246849
44	80.6	5.8	843	9	CG254385
45	80.6	5.8	852	8	CC385881

ALIGNMENTS

RESULT 1
LOCUS CC656933 963 bp DNA linear GSS 19-JUN-2003
DEFINITION OGMDO20TM ZM 0.7-1.5 KB Zea mays genomic clone ZMMBMA0554D15,
genomic survey sequence.
ACCESSION CC656933
VERSION CC656933.1 GI:32060225
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 963)

REFERENCE
AUTHORS Resnick,A., Fraser,C.M., Buddiman,M.A., Bedell,J.A., Rohlfing,T.,
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Uteerback,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
TITLE Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Other GSSs: OGMDO20TV
Contact: Cathy Whitelaw
TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.

FEATURES
source location/Qualifiers
1..963

/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZM_0.7-1.5_KB"
/note="Vector: pBSCK-; Site:1: HincII, 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match 68.3%; Score 952; DB 9; Length 963;

Best Local Similarity 99.9%; Pred. No. 3.8e-191;

Matches 963; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

404 ACCAGCATGTCTAGTCTCACTATTTGGCACTACCAAGCAAGATTTAAAAAATTAACC 463
1 ACCAGCATGTCTAGTCTCACTATTTGGCACTACCAAGCAAGATTTAAAAAATTAACC 60

```
OY 464 AAAGTAAGTAATCCAGTCGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAACC 523
    |||
DB 61 AAGTAAGTAATCCAGTCGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAACC 120
OY 524 ACGATCCTCTTAAACAAAGCATATTTGCAAGAGCAAAATATGTTACAGTTTACAA 583
    |||
DB 121 ACGATCCTCTTAAACAAAGCATATTTGCAAGAGCAAAATATGTTACAGTTTACAA 180
OY 584 CATCTAAGAGCGCAAAATATATCGAAAGGTATGACGCTTCAGATTTTCTTTTC 643
    |||
DB 181 CATCTAAGAGCGCAAAATATATCGAAAGGTATGACGCTTCAGATTTTCTTTTC 240
OY 644 ATCTGTATTTGTTGTTATGTTTATATACATTTTCTCTCTCAATAGAGATTT 703
    |||
DB 241 ATCTGTATTTGTTGTTATGTTTATATACATTTTCTCTCTCAATAGAGATTT 300
OY 704 TCTTCGATTTTAAATGACTATTAAGTCAATTTTATATAAGACAGCATGTCGTAG 763
    |||
DB 301 TCTTCGATTTTAAATGACTATTAAGTCAATTTTATATAAGACAGCATGTCGTAG 360
OY 764 ATTCTGTTCAAAATCTTCTGATTTTAAAGAGTGTGCAACCTGTTCTTT 823
    |||
DB 361 ATTCTGTTCAAAATCTTCTGATTTTAAAGAGTGTGCAACCTGTTCTTT 420
OY 824 CAAGAATTTTGAATTTTCAAAAAATATGTTTATTTTCTTATATAAATAGAAAC 883
    |||
DB 421 CAAGAATTTTGAATTTTCAAAAAATATGTTTATTTTCTTATATAAATAGAAAC 480
OY 884 ACTTGAAGAAATAGAGTGGCAGAGTACCCCTGAGATGTTTCCCAATTAATCAATCA 943
    |||
DB 481 ACTTGAAGAAATAGAGTGGCAGAGTACCCCTGAGATGTTTCCCAATTAATCAATCA 540
OY 944 CTGTGATTAATTTTGGCCAGGCCCATTAATTTTAAACCAACTGAAATCGAGCA 1003
    |||
DB 541 CTGTGATTAATTTTGGCCAGGCCCATTAATTTTAAACCAACTGAAATCGAGCA 600
OY 1004 AACCAATCTGAGCTATTTCTCTAGATTAGTAAAGAGAGAGAGAGAGAGAAATCAG 1063
    |||
DB 601 AACCAATCTGAGCTATTTCTCTAGATTAGTAAAGAGAGAGAGAGAGAAATCAG 660
OY 1064 TTTTAAGTATGTCCTGAGATGTCGGTTTGGCAAGATAGCCACCGTAATCATAGCT 1123
    |||
DB 661 TTTTAAGTATGTCCTGAGATGTCGGTTTGGCAAGATAGCCACCGTAATCATAGCT 720
OY 1124 CAAAGTGCTAGTCAAGTTCGGCAGCTTCCTGTCACTCACTCACTGCACTACTATATGC 1183
    |||
DB 721 CAAAGTGCTAGTCAAGTTCGGCAGCTTCCTGTCACTCACTCACTGCACTACTATATGC 780
OY 1184 TTGTTCAACCGTTCGTCTGTTCATGTCGAAGCCCTTGCTATTCTGAACCAAGAGAT 1243
    |||
DB 781 TTGTTCAACCGTTCGTCTGTTCATGTCGAAGCCCTTGCTATTCTGAACCAAGAGAT 840
OY 1244 ACCTACTCCCAAAACAATTCATCTTACTCATGCAACTTCACGAAACACGACATATGTT 1303
    |||
DB 841 ACCTACTCCCAAAACAATTCATCTTACTCATGCAACTTCACGAAACACGACATATGTT 900
OY 1304 TCTTGAACCAATTCATTAAGATCAACAGCTAAGGTTCTCCCGTACTTCCCTCTCT 1363
    |||
DB 901 TCTTGAAC-ATCCATTAAGATCAACAGCTAAGGTTCTCCCGTACTTCCCTCTCT 959
OY 1364 CCTC 1367
    |||
DB 960 CCTC 963
```

```
RESULT 2
CG224225
LOCUS CG224225 915 bp DNA linear GSS 22-AUG-2003
DEFINITION OGIAG08YV_ZM_0_7_1_5_KB Zea mays genomic clone ZMMBMA0716B15,
genomic survey sequence.
ACCESSION CG224225
VERSION CG224225.1 GI:34124113
```

```
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE 1 (bases 1 to 915)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Uteerback,T.,
Reinick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nundberg,A., Rodbins,D. and Lakey,N.
TITLE Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Other GSSs: OGIAG08TH
Contact: Cathy Whitelaw
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@cigr.org
Seq primer: TP
Classes: sheared ends.
FEATURES
    source
        location/Qualifiers
            1..915
            /organism="Zea mays"
            /mol_type="genomic DNA"
            /strain="B73"
            /db_xref="taxon:4577"
            /clone="ZMMBMA0716B15"
            /clone_11b="ZM_0_7_1_5_KB"
            /note="vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
            methylation filtered genomic DNA library"
ORIGIN
Query Match 54.5%; Score 760.4; DB 9; Length 915;
Best Local Similarity 99.7%; Pred. No. 1.5e-150;
Matches 772; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
OY 621 TGACGTCAGATTTTCTTTTTCATCTGTTATTTGTTATGTTTAAATACATTTT 680
    |||
DB 1 TGACGTCAGATTTTCTTTTTC- TTCTTGTTATTTGTTATGTTTAAATACATTTT 59
OY 681 CTTCCTTACAATAGATGATTTTCTCCGATTTTATATAATGACTATAAGTCAATTTT 740
    |||
DB 60 CTTCCTTACAATAGATGATTTTCTCCGATTTTATATAATGACTATAAGTCAATTTT 119
OY 741 ATATAAGACACGCAATGTCGATTCCTGTTCAAAAATCTTTCGATTTTAAAGAC 800
    |||
DB 120 ATATAAGACACGCAATGTCGATTCCTGTTCAAAAATCTTTCGATTTTAAAGAC 179
OY 801 TAGTTGGCAACCCGTCTTCTTCAAGAAATTTGATTTTCAAAAAATTAGTTTAT 860
    |||
DB 180 TAGTTGGCAACCCGTCTTCTTCAAGAAATTTGATTTTCAAAAAATTAGTTTAT 239
OY 861 TTTCTCTTATTAATAGAAAAACCTTAGAAAAATAGATTGCCAGACTAGCCCTAGAT 920
    |||
DB 240 TTTCTCTTATTAATAGAAAAACCTTAGAAAAATAGATTGCCAGACTAGCCCTAGAT 299
OY 921 GTTTTCCCAATTAATCAATCACTGTATTAATTTTGGCCAGCCCATTAATTTT 980
    |||
DB 300 GTTTTCCCAATTAATCAATCACTGTATTAATTTTGGCCAGCCCATTAATTTT 359
OY 981 AAACGAACTGAATTCGAGGAAACCAATCTGAGTATTTCTAGATTAGTAAAG 1040
    |||
DB 360 AAACGAACTGAATTCGAGGAAACCAATCTGAGTATTTCTAGATTAGTAAAG 419
OY 1041 GGAGAGAGAGAGAAATCAATTAAGTCAATTTGCTGAGATGTCGGTTTGCA 1100
    |||
DB 420 GGAGAGAGAGAGAAATCAATTAAGTCAATTTGCTGAGATGTCGGTTTGCA 479
OY 1101 CGATAGCCACCGTAATCAATAGTCAATAGTGCCTACGTCAGGTTGGCAGCTCTGTGTC 1160
    |||
DB 480 CGATAGCCACCGTAATCAATAGTCAATAGTGCCTACGTCAGGTTGGCAGCTCTGTGTC 539
```

QY 1161 ATCTACATGGCATACTACATGCTTGTTCACACGTTGCTGTGTTCCATGCTCCAGCCT 1220
|||||
Db 540 ATCTACATGGCATACTACATGCTTGTTCACACGTTGCTGTGTTCCATGCTCCAGCCT 599
|||||
QY 1221 TGCTATTCTGAACCAAGAGATACCTACTCTCCAAACATTCATCTTACTGACCACTT 1280
|||||
Db 600 TGCTATTCTGAACCAAGAGATACCTACTCTCCAAACATTCATCTTACTGACCACTT 659
|||||
QY 1281 CCATGCAACACGACATATGTTCTCTGAACCAATTCATTAAGATACACAACAGCTAGCG 1340
|||||
Db 660 CCATGCAACACGACATATGTTCTCTGAACCAATTCATTAAGATACACAACAGCTAGCG 719
|||||
QY 1341 TTCTCCGCTAGCTGCTCTCTCTGCGATCTTTTTCGTCACACCATG 1394
|||||
Db 720 TTCTCCGCTAGCTGCTCTCTCTCTGCGATCTTTTTCGTCACACCATG 773
|||||
RESULT 3
CC656939/c 687 bp DNA linear GSS 19-JUN-2003
DEFINITION OGMDQ20TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMA0554D15,
genomic survey sequence.
ACCESSION CC656939
VERSION CC656939.1 GI:32060231
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 687)
REFERENCE Whitefaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
OTHER_GSS: OGMDQ20TM
COMMENT Other: Cathy Whitefaw
Contact: Cathy Whitefaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitefaw@tigr.org
Seq primer: TF
Class: sheared ends.
FEATURES
source location/Qualifiers
1..687
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0554D15"
/clone_1lb="ZM_0.7_1.5_KB"
/note="Vector: pBCK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"
ORIGIN
Query Match 36.4%; Score 501.2; DB 9; Length 687;
Best Local Similarity 99.4%; Pred. No. 1,2e-95;
Matches 503; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 889 GAAAAATAGAGTGGCCAGACTAGCCCTAGAGATGTTTCCCAATTAATTACATCACTGTG 948
|||||
Db 687 GAAAAATAGAGTGGCCAGACTAGCCCTAGAGATGTTTCCCAATTAATTACATCACTGTG 628
|||||
QY 949 TATAATATTGTCGACGCCCATTAATTATTAACGGAATCGAATTCAGAGAAACCA 1008
|||||
Db 627 TATAATATTGTCGACGCCCATTAATTATTAACGGAATTCAGAGAAACCA 568
|||||
QY 1009 AATCTGACTATTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAGAAATCAGTTT 1068
|||||
Db 567 AATCTGACTATTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAGAAATCAGTTT 508
|||||

QY 1069 AGTCATTGTCCTGAGATGNGCGTTTGGCAACGATAGCCGTAATCATAGCTATAG 1128
|||||
Db 507 AGTCATTGTCCTGAGATGNGCGTTTGGCAACGATAGCCGTAATCATAGCTATAG 448
|||||
QY 1129 GTGCTACGTCAGAGTTCGAGAGCTCTGTGTCATCTCAATGAGCATACATGCTTGT 1188
|||||
Db 447 GTGCTACGTCAGAGTTCGAGAGCTCTGTGTCATCTCAATGAGCATACATGCTTGT 388
|||||
QY 1189 CAACGCTTGTCTTGTTCATGTCGACGCTTTCCTATTCTGAACCAAGAGATACCTA 1248
|||||
Db 387 CAACGCTTGTCTTGTTCATGTCGACGCTTTCCTATTCTGAACCAAGAGATACCTA 328
|||||
QY 1249 CTCCTCAACCAATCCATCTTATCTGACCACTTCGACCAACGACATATCTTCG 1308
|||||
Db 327 CTCCTCAACCAATCCATCTTATCTGACCACTTCGACCAACGACATATCTTCG 268
|||||
QY 1309 AACCAATCCATTAAGATACAAACAGCTAGCTTCCGCTAGCTTCTCTCTCT 1368
|||||
Db 267 AACCAATCCATTAAGATACAAACAGCTAGCTTCTCTCTCTCTCTCTCTCT 208
|||||
QY 1369 GCCGATCTTTTGTGTCACACCATG 1394
|||||
Db 207 GCCGATCTTTTGTGTCACACCATG 182
|||||
RESULT 4
CG252571/c 715 bp DNA linear GSS 25-AUG-2003
DEFINITION OG4B8057C ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMA0809B10,
genomic survey sequence.
ACCESSION CG252571
VERSION CG252571.1 GI:34154661
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 715)
REFERENCE Whitefaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
COMMENT Other: Cathy Whitefaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitefaw@tigr.org
Seq primer: TF
Class: sheared ends.
FEATURES
source location/Qualifiers
1..715
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0809B10"
/clone_1lb="ZM_0.7_1.5_KB"
/note="Vector: pBCK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"
ORIGIN
Query Match 6.9%; Score 96; DB 9; Length 715;
Best Local Similarity 80.1%; Pred. No. 8.4e-10;
Matches 125; Conservative 0; Mismatches 30; Indels 1; Gaps 1;
QY 135 TCTCCACCTCGGCTGCACTTTCTTGATGTCGGTGTCCATCTGACGAGGCCCA 194
|||||
Db 356 TCAACACCGCTCTTGTGTCGATGTCGGGCGGTTCCCGCTCATGAGGCCCCA 297
|||||
QY 195 TCAACACCTCTTTCGGAACCACTCAAGGCTTTTCGATGCGCCCAAGAGATCGG 254
|||||

Db 296 TAAGGACCTGTGAGAAACCATTAAGGCTTTGGATGGCCACGGAGCAT-CGGGG 238
 QY 255 TCCTGGATCCAGGGGATATATGTCCTCCCAATC 290
 Db 237 TCCTGGATCCAGGGGATATATGTCCTCCCAATC 202

RESULT 5
 CL235046 967 bp DNA linear GSS 15-JAN-2004
 LOCUS ZMBBB0575001.r ZMBBB (HindIII) Zea mays genomic clone
 DEFINITION ZMBBB0575001 3', genomic survey sequence.
 ACCESSION CL235046
 VERSION CL235046.1 GI:40891729
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 967)
 Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C.,
 Zohovetz,V., Fuks,G., Yu,Y., Wing,R. and Messing,J.
 Sequencing of the maize genome at PGIR (2003c)
 Unpublished (2003)
 Contact: Bharti,A.K.
 Dr.Joachim Messing's lab
 The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
 University
 190 Frelinghuysen Road, Piscataway, NJ 08854, USA
 Tel: 732 445 3801
 Fax: 732 445 5735
 Email: bharti@waksman.rutgers.edu
 Seq primer: SP6
 Class: BAC ends
 High quality sequence start: 64.
 Location/Qualifiers
 1. 967
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /culivar="B73"
 /db_xref="taxon:4577"
 /clone="ZMBBB0575001"
 /lab_host="E. coli DH10B"
 /clone_lib="ZMBBB (HindIII)"
 /note="Vector: pCUG1; Site_1: HindIII; Site_2: HindIII"

ORIGIN

Query Match 6.6%; Score 92; DB 9; Length 967;
 Best Local Similarity 78.6%; Pred. No. 6e-09;
 Matches 110; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 150 TGCACATTTCTTGGATGTCGGTTCCTGATGACCGAGCCCATCAACACCTTTGG 209
 Db 827 TGGCCCTTTCTTGATACCGCGCGTTCCCTGTGATGAGGCCCATCCAGACCTTTTGG 886
 QY 210 GACACCATCAAGGGCTTTCGATGCGCCAGAGAGATGCGGTGCGTGTGATCCAG 269
 Db 887 GTGGCCCATCAAGTGCCTTTCGATGTCACCGAGATGATCCGGGCCGTGTGACCTGG 946
 QY 270 GGATATATGTCCCCCAAT 289
 Db 947 GGATATCTGTCCCCCAATAGT 966

RESULT 6
 CG048704 814 bp DNA linear GSS 19-AUG-2003
 LOCUS PUILO19TB.ZM.0.6.1.0 KB Zea mays genomic clone ZMBBTA0611C13,
 DEFINITION genomic survey sequence.
 ACCESSION CG048704
 VERSION CG048704.1 GI:33920884

KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 814)
 Whitefaw,C.A., Quackenbush,J., Van Aken,S., Uteerback,T.,
 Renick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
 Benner,J.
 Maize Genomics Consortium
 Unpublished (2003)
 Other GSSs: PUILO19TB
 Contact: Cathy Whitefaw
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whitefaw@tigr.org
 Seq primer: TR
 Class: sheared ends.
 Location/Qualifiers
 1. 814
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone="ZMBBTA0611C13"
 /clone_lib="ZM_0.6.1.0_KB"
 /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
 cot selected genomic DNA library"

ORIGIN

Query Match 6.3%; Score 88; DB 9; Length 814;
 Best Local Similarity 62.7%; Pred. No. 4.2e-08;
 Matches 153; Conservative 0; Mismatches 90; Indels 1; Gaps 1;

QY 713 TTTATAAATGACTTAAAGTCATTTTATATPAAGACACGCGATGCTGATTCGCTT 772
 Db 487 TTTAAGACAGCTGTGTATAGATGAGAAAGATATAGAAATATCTTTAGGAATGTTA 546
 QY 773 CAAAAATCTTCTGATTTTAAAGAGCTAGTTGGCAACCTGTCTTCAAGAATT 832
 Db 547 TAAAGACACAGAAATTTCTTAAGAGCTAGTTGTATCCCATTTTCAAGAATT 606
 QY 833 TTGATTTTTCAAA-AAATATGTTATTTCTCTTATATAAATGAAAACATTAGA 891
 Db 607 TTCATTCACAAAGGAAATATGTTATTTCCCTTTGTAAATGTAATCAGTCGGA 666
 QY 892 AAATAGAGTTCACACTAGCCCTAGAAATGTTTCCCAATAATTAACAATCACTGTAT 951
 Db 667 AAATAGAGTTCACAACTAGCCCTAAGATATATCCGATATATCCCATATCTCTCT 726
 QY 952 AATT 955
 Db 727 ATTT 730

RESULT 7
 CG414922/c 754 bp DNA linear GSS 08-SEP-2003
 LOCUS ZMBBB0290L09.r ZMBBB Zea mays genomic clone ZMBBB0290L09 3',
 DEFINITION genomic survey sequence.
 ACCESSION CG414922
 VERSION CG414922.1 GI:34505144
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 754)
 Yu,Y., Kim,H.R., Hatfield,J., Soderlund,C., Bharti,A.K., Messing,J.

TITLE and Wing, R.
JOURNAL Sequencing of the maize genome
COMMENT Unpublished (2003)
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 440A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: <http://genome.arizona.edu>
PCR Primers
FORWARD: T7
BACKWARD: M13r
Plate: 0290 row: L column: 09
Seq primer: M13r
Class: BAC ends.

FEATURES
source location/Qualifiers

1..754
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ZMMBB0290L09"
/lab_host="DH108"
/clone_lib="ZMMBBb"
/note="Vector: pBel0BAC11, Site_1: HindIII, Site_2:
HindIII; Zea mays L. sep. mays"

ORIGIN

Query Match 6.3%; Score 87.8; DB 9; Length 754;
Best Local Similarity 71.2%; Pred. No. 4.6e-08;

Matches 116; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 779 TCTTTCGATTTTAAAGAGCTAGTTGGCAACCGTCTTCTTCAAGATTGATT 838
DB 680 TCTAAATTATTCCTAAAGCTAGTTGAAAGCAGATTTTCCAGATTTTAAAT 621
QY 839 TTTTCAAAAAAATTAGTTATTTCTCTTATTAATAAGAAAACACTAGAAAATAGA 898
DB 620 TTTCTAAATAAATAATATTTCTTGAATAAATAGAAATCAGATGAAAAATAA 561
QY 899 GTTGGCAGACTAGCCCTAGAAATGTTTCCCAATTAATTAAT 941
DB 560 GTTCTTAACCTAACTAAATAAATTTCTTCGTCATTACAGT 518

RESULT 8

CC439901 950 bp DNA linear GSS 20-MAY-2003
LOCUS PUHRV15TB_ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa519D06,
DEFINITION genomic survey sequence.
ACCESSION CC439901
VERSION CC439901.1 GI:30940477
KEYWORDS GSS.

SOURCE

ORGANISM Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 950)
Whiteley, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and
Bennetzen, J.

REFERENCE

AUTHORS
TITLE JOURNAL
COMMENT Unpublished (2003)
Other_GSSs: PUHRV15TD
Contact: Cathy Whiteley
TIGR

712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteley@cigr.org

Seq primer: TR
Class: sheared ends.
Location/Qualifiers
1..950
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTa519D06"
/note="Vector: PCR4-TOPO, Site_1: EcoRI; 0.6-1.0 kb high
Cot selected genomic DNA library"

ORIGIN

Query Match 6.2%; Score 86.8; DB 8; Length 950;
Best Local Similarity 72.4%; Pred. No. 7.6e-08;
Matches 126; Conservative 0; Mismatches 47; Indels 1; Gaps 1;

QY 787 ATTTTAAAGAGCTAGTTGGCAACCGTCTTCAAGATTGATTTTCAAA 846
DB 241 ATCTTCTAATGACTAGTTGGCAACCTTAATTT-TCTCAGGGGATTTTATTTTCAAT 183
QY 847 AAAAATTAGTTATTTCTCTTATTAATAAGAAAACCTAGAAAATAGAGTCCAG 906
DB 182 GAAAATTAGTTATTTCTCTTGAAGAAATATTAATCATTGAGAAAATATAGTTCCAA 123
QY 907 ACTAGCCCTAGAAATGTTTCCCAATTAATTAATCACTGTGTATTAATTTG 960
DB 122 ACTAGCCCTAAGATTAATTAATAATAGAAAATAATGTCTATTTTCAATG 69

RESULT 9

CC384247 652 bp DNA linear GSS 19-MAY-2003
LOCUS PUHOC67TB_ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa495L13,
DEFINITION genomic survey sequence.
ACCESSION CC384247
VERSION CC384247.1 GI:30864143
KEYWORDS GSS.

SOURCE

ORGANISM Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 652)
Whiteley, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and
Bennetzen, J.

REFERENCE

AUTHORS
TITLE JOURNAL
COMMENT Unpublished (2003)
Other_GSSs: PUHOC67TD
Contact: Cathy Whiteley
TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteley@cigr.org

Seq primer: TR
Class: sheared ends.

FEATURES

source location/Qualifiers
1..652
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTa495L13"
/clone_lib="ZM_0.6_1.0_KB"
/note="Vector: PCR4-TOPO, Site_1: EcoRI; 0.6-1.0 kb high
Cot selected genomic DNA library"

ORIGIN

Query Match 6.2%; Score 86.6; DB 8; Length 652;
Best Local Similarity 63.9%; Pred. No. 8.2e-08;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using bw model

Run on: September 15, 2005, 08:25:31 ; Search time 526.823 Seconds
(without alignment)
14440.280 Million cell updates/sec

Title: US-10-713-381-1_COPY_1155_1311

Perfect score: 157
Sequence: 1 cgtgcattccatccatgcat.....cgcacatcgtcttcctgaac 157

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Genemb1:*
1: gb_ba:*
2: gb_hgi:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sbs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	157	100.0	255	6	AX224402 Sequence
2	157	100.0	1394	6	AX224394 Sequence
3	157	100.0	1394	6	AX224395 Sequence
4	157	100.0	1394	6	BD062176 Male t188
5	157	100.0	1394	6	BD062177 Male t188
6	157	100.0	3343	8	AF360356 Zea mays
7	146	93.0	158	6	AX224396 Sequence
8	54	34.4	137327	8	AC135206 Oryza sat
9	50	31.8	50	6	AX224398 Sequence
10	40	25.5	40	6	AX224399 Sequence
11	34.8	22.2	174806	2	BX950854 Dario rer
12	34.6	22.0	110137	9	AC104134 Homo sapi
13	34.4	21.9	214966	2	CR847532 Dario rer
14	33.8	21.5	141810	2	CR847532 Dario rer
15	33.8	21.5	214178	5	BX640577 Zebrafish
16	33.2	21.1	117764	9	AC0911290 Homo sapi
17	33.2	21.1	162481	2	AC161334 Pan trogl
18	32.8	20.9	82642	9	AC103745 Homo sapi

C	20	32.8	20.9	93610	9	AC006602	Homo sapi
C	21	32.8	20.9	130031	2	AC135254	Homo sapi
C	22	32.8	20.9	148598	9	HSB951C14	Human DNA
C	23	32.8	20.9	158316	9	AC135350	Homo sapi
C	24	32.8	20.9	169771	9	AC087485	Homo sapi
C	25	32.8	20.9	169667	2	AC019229	Homo sapi
C	26	32.8	20.9	193569	2	AC149044	Pan trogl
C	27	32.6	20.8	53688	2	AC099973	Mus muscu
C	28	32.6	20.8	146800	5	AL953855	Zebrafish
C	29	32.6	20.8	165845	2	AC127012	Rattus no
C	30	32.6	20.8	204190	2	AC147308	Pan trogl
C	31	32.6	20.8	204190	2	AC140696	Rattus no
C	32	32.6	20.8	252534	2	AC097405	Rattus no
C	33	32.4	20.6	2070	10	BC020027	Mus muscu
C	34	32.4	20.6	2096	10	BC038066	Mus muscu
C	35	32.4	20.6	2433	10	BC055469	Mus muscu
C	36	32.4	20.6	110000	1	AE017197_05	Continuation (6 of
C	37	32.4	20.6	137832	2	AC104550	Mus muscu
C	38	32.4	20.6	168889	10	AL731779	Mouse DNA
C	39	32.4	20.6	182847	2	AC148229	Gallus ga
C	40	32.4	20.6	221338	5	AC147437	Gallus ga
C	41	32.4	20.6	262640	2	AC139057	Mus muscu
C	42	32.2	20.5	139173	9	AC009140	Homo sapi
C	43	32.2	20.5	150116	2	AC034132	Homo sapi
C	44	32.2	20.5	189323	2	AC150620	Callithri
C	45	32.2	20.5	205954	9	AC008546	Homo sapi

ALIGNMENTS

RESULT 1	AX224402	255 bp	DNA	linear	PAT 10-SEP-2001
LOCUS	AX224402	Sequence 9 from Patent WO0160997.			
DEFINITION	AX224402				
ACCESSION	AX224402				
VERSION	AX224402.1	GI:15554644			
KEYWORDS					
SOURCE					
ORGANISM					
	Zea mays				
	Zea mays				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.				
REFERENCE					
AUTHORS	Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.				
TITLE	Male tissue-preferred regulatory region and method of using same				
JOURNAL	Patent: WO 0160997-A 9 23-AUG-2001;				
PIONEER	HI-BRED INTERNATIONAL, INC. (US)				
FEATURES					
source	1..255				
	/organism="Zea mays"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:4577"				
ORIGIN					
Query Match	100.0%;	Score 157;	DB 6;	Length 255;	
Best Local Similarity	100.0%;	Pred. No. 1.8e-40;			
Matches 157;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	1	CGTGCATTCACATGCGATACATGCTGTTCACCGTTCGTTTCATGATGCG	60		
DB	15	CGTGCATTCACATGCGATACATGCTGTTCACCGTTCGTTTCATGATGCG	74		
QY	61	AAGCGTTCGCTATTTCGAACGAAGGATACCTACCTCCAAACATTCATTCATG	120		
DB	75	AAGCGTTCGCTATTTCGAACGAAGGATACCTACCTCCAAACATTCATTCATG	134		
QY	121	CAACTTCATGCAACGCAATATGTTTCTGAAAC	157		
DB	135	CAACTTCATGCAACGCAATATGTTTCTGAAAC	171		
RESULT 2					

AX224394
LOCUS AX224394 1394 bp DNA linear PAT 10-SEP-2001
DEFINITION Sequence 1 from Patent WO0160997.
ACCESSION AX224394
VERSION AX224394.1 GI:15554636
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
AUTHORS Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
TITLE Male tissue-preferred regulatory region and method of using same
JOURNAL Patent: WO 0160997-A 1 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
source location/Qualifiers
1..1394
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN
Query Match 100.0%; Score 157; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 1.9e-40;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGTGATCTGCATGACATGACATGCTGTTTCAACCGTTCCTGTTCCATGCTGC 60
DB 1155 CGGTGATCTGCATGACATGACATGCTGTTTCAACCGTTCCTGTTCCATGCTGC 1214

QY 61 AAGCTTGCTATTTGGAACCAAGAGATGACTACTCTCCCAACATCCATCTTACTCATG 120
DB 1215 AAGCTTGCTATTTGGAACCAAGAGATGACTACTCTCCCAACATCCATCTTACTCATG 1274

QY 121 CAATTCATGCAACACGACATATGTTTCTGGAAC 157
DB 1275 CAATTCATGCAACACGACATATGTTTCTGGAAC 1311

RESULT 3
AX224395
LOCUS AX224395 1394 bp DNA linear PAT 10-SEP-2001
DEFINITION Sequence 2 from Patent WO0160997.
ACCESSION AX224395
VERSION AX224395.1 GI:15554637
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
AUTHORS Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
TITLE Male tissue-preferred regulatory region and method of using same
JOURNAL Patent: WO 0160997-A 2 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
source location/Qualifiers
1..1394
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN
Query Match 100.0%; Score 157; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 1.9e-40;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGTGATCTGCATGACATGACATGCTGTTTCAACCGTTCCTGTTCCATGCTGC 60
DB 1155 CGGTGATCTGCATGACATGACATGCTGTTTCAACCGTTCCTGTTCCATGCTGC 1214

QY 61 AAGCTTGCTATTTGGAACCAAGAGATGACTACTCTCCCAACATCCATCTTACTCATG 120
DB 1215 AAGCTTGCTATTTGGAACCAAGAGATGACTACTCTCCCAACATCCATCTTACTCATG 1274

QY 121 CAATTCATGCAACACGACATATGTTTCTGGAAC 157
DB 1275 CAATTCATGCAACACGACATATGTTTCTGGAAC 1311

DB 1215 AAGCTTGCTATTTGGAACCAAGAGATGACTACTCTCCCAACATCCATCTTACTCATG 1274

QY 121 CAATTCATGCAACACGACATATGTTTCTGGAAC 157
DB 1275 CAATTCATGCAACACGACATATGTTTCTGGAAC 1311

RESULT 4
BD062176
LOCUS BD062176 1394 bp DNA linear PAT 27-AUG-2002
DEFINITION Male tissue-preferred regulatory region and method of using same.
ACCESSION BD062176
VERSION BD062176.1 GI:22607781
KEYWORDS JP 2001520523-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
AUTHORS Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G.A. and Kendall,T.L.
TITLE Male tissue-preferred regulatory region and method of using same
JOURNAL Patent: JP 2001520523-A 1 30-OCT-2001;
PIONEER HI BRED INTERNATIONAL INC
COMMENT PN JP 2001520523-A/1
PD 30-OCT-2001
PF 19-JUN-1998 JP 1998504910
PR 23-JUN-1997 US 08/880499
PI MARC C ALBERTSEN,TIMOTHY W FOX,CARL W GARNAAT,GARY A HUFFMAN,
PI TIMMY L KENDALL
PC C12N15/82,C12N15/29,C12N9/24,C12N9/22,C12N9/10,C12N9/00 PC
PC C07K14/34,C12Q1/68,
PC A01H5/00
CC Strandedness: Single;
CC Topology: Linear;
FH Key location/Qualifiers
1..1394
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 157; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 1.9e-40;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGTGATCTGCATGACATGACATGCTGTTTCAACCGTTCCTGTTCCATGCTGC 60
DB 1155 CGGTGATCTGCATGACATGACATGCTGTTTCAACCGTTCCTGTTCCATGCTGC 1214

QY 61 AAGCTTGCTATTTGGAACCAAGAGATGACTACTCTCCCAACATCCATCTTACTCATG 120
DB 1215 AAGCTTGCTATTTGGAACCAAGAGATGACTACTCTCCCAACATCCATCTTACTCATG 1274

QY 121 CAATTCATGCAACACGACATATGTTTCTGGAAC 157
DB 1275 CAATTCATGCAACACGACATATGTTTCTGGAAC 1311

RESULT 5
BD062177
LOCUS BD062177 1394 bp DNA linear PAT 27-AUG-2002
DEFINITION Male tissue-preferred regulatory region and method of using same.
ACCESSION BD062177
VERSION BD062177.1 GI:22607782
KEYWORDS JP 2001520523-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
1 (bases 1 to 1394)

AUTHORS Albertsen, M.C., Fox, T.W., Garnaat, C.W., Huffman, G.A. and Kendall, T.L.

TITLE Male tissue-preferred regulatory region and method of using same

JOURNAL Patent: JP 2001520523-A 2 30-OCT-2001;
PIONEER HI-BRED INTERNATIONAL INC

COMMENT PN JP 2001520523-A/2
PD 30-OCT-2001
PF 19-JUN-1998 JP 1999504910
PR 23-JUN-1997 US 08/880499
PI MARC C ALBERTSEN, TIMOTHY W FOX, CARL W GARNAA, GARY A HUFFMAN,
TIMMY L KENDALL
PC C12N15/82, C12N15/29, C12N9/24, C12N9/22, C12N9/10, C12N9/00 PC
C07K14/34, C12Q1/68,
PC A01H5/00
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers.

FEATURES source 1..1394
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 157; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 1.9e-40;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTATCTACATGGCATCTACATGCTGTTCAACGCTTGCTGTTCCATGCTCC 60
DB 1155 CGTGTATCTACATGGCATCTACATGCTGTTCAACGCTTGCTGTTCCATGCTCC 1214

QY 61 AAGCCTTGCTATTCTGAAACCAAGAGATACCTACTCCAAACATCATCTTACTGATG 120
DB 1215 AAGCCTTGCTATTCTGAAACCAAGAGATACCTACTCCAAACATCATCTTACTGATG 1274

QY 121 CAACCTTCATGCAACACGACATATGTTCTCGAAC 157
DB 1275 CAACCTTCATGCAACACGACATATGTTCTCGAAC 1311

RESULT 6
AF360356 3343 bp DNA linear PLN 12-MAY-2001
LOCUS AF360356 Zea mays male fertility protein (M645) gene, complete cds.
DEFINITION
ACCESSION AF360356
VERSION AF360356.1 GI:14028756
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 3343)
Fox, T.W., Trimmell, M.R. and Albertsen, M.C.
Cloning of M645, a gene required for male fertility from Zea mays
unpublished
2 (bases 1 to 3343)
Fox, T.W., Trimmell, M.R. and Albertsen, M.C.
Direct Submission
Submitted (13-MAR-2001) Trait and Technology Development, Pioneer
Hi-Bred Intl. Inc., 7300 N.W. 62nd Ave., P.O. Box 1004, Johnston,
IA 50131-1004, USA
Location/Qualifiers
source 1..3343
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/chromosome="9L"
<1392..>2942
/gene="M645"
join(<1392..1768,1898..2182,2280..2447,2534..>2942)

CDS

/gene="M645"
/product="male fertility protein"
join(1392..1768,1898..2182,2280..2447,2534..2942)
/gene="M645"
/codon_start=1
/product="male fertility protein"
/protein_id="AAK52489.1"
/db_xref="GI:14028757"
/translation="MEKRLQWRKRGDIQVYPHLFFAALALALVADPFGSLPAEAV
DYRPVGHLEAPYGEVWSPNDASRLRGRLFYGVGFPESIEFDLOGRPYAGLA
DGRVVMGEBAEMETFAVMPNDMSRVCANGVNSTTRQHEKEEGPGLGRFGE
TGEIYVADAVYGLMVGSGGVASSVAREADDPIRFANDLVHNRGSYFPDTSRY
SKRDHNTLLIEBEGRLRYRPERSGVYVYAKGVFPNGVQISEDHQFLISSETNC
RIMRWLBSGRAGEVEFANLPFGPDNVNSNRGQFVAIDCCTPDAQEVFAKPMLR
TLVFKPLSLKVLTWKARMTVALLDGEGRVVEVLDGRHEVWKLVSVEAVEGRK
LWIGVAHNHVIATIPYPLD"

ORIGIN

Query Match 100.0%; Score 157; DB 8; Length 3343;
Best Local Similarity 100.0%; Pred. No. 1.9e-40;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTATCTACATGGCATCTACATGCTGTTCAACGCTTGCTGTTCCATGCTCC 60
DB 1155 CGTGTATCTACATGGCATCTACATGCTGTTCAACGCTTGCTGTTCCATGCTCC 1214

QY 61 AAGCCTTGCTATTCTGAAACCAAGAGATACCTACTCCAAACATCATCTTACTGATG 120
DB 1215 AAGCCTTGCTATTCTGAAACCAAGAGATACCTACTCCAAACATCATCTTACTGATG 1274

QY 121 CAACCTTCATGCAACACGACATATGTTCTCGAAC 157
DB 1275 CAACCTTCATGCAACACGACATATGTTCTCGAAC 1311

RESULT 7
AX224396 158 bp DNA linear PAT 10-SEP-2001
LOCUS AX224396 Sequence 3 from Patent WO0160997.
DEFINITION
ACCESSION AX224396
VERSION AX224396.1 GI:15554638
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
1
Albertsen, M.C., Fox, T.W., Garnaat, C.W., Huffman, G. and Kendall, T.L.
Male tissue-preferred regulatory region and method of using same
Patent: WO 0160997-A 3 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
Location/Qualifiers
source 1..158
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN

Query Match 93.0%; Score 146; DB 6; Length 158;
Best Local Similarity 99.4%; Pred. No. 7.2e-37;
Matches 157; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 CGTGTATCTACATGGCATCTACATGCTGTTCAACGCTTGCTGTTCCATGCTCC 59
DB 1 CGTGTATCTACATGGCATCTACATGCTGTTCAACGCTTGCTGTTCCATGCTCC 60

QY 60 CAAGCTTGCTATTCTGAAACCAAGAGATACCTACTCCAAACATCATCTTACTGAT 119
DB 61 CAAGCTTGCTATTCTGAAACCAAGAGATACCTACTCCAAACATCATCTTACTGAT 120

QY 120 GCAACTTCATGCAACACGACATATGTTCTCGAAC 157
DB 1275 GCAACTTCATGCAACACGACATATGTTCTCGAAC 1311

Db 121 GCACTTCATGCAACAGCACATATGTTTCTGAAC 158

RESULT 8
AC135206/c 137327 bp DNA linear PLN 16-APR-2003
LOCUS Oryza sativa (japonica cultivar-group) chromosome 3 clone
DEFINITION OJ1041F02, complete sequence.
AC135206
AC135206.3 GI:27596977
KEYWORDS HTG.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Eriocaridaceae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 137327)
Wing,R.A., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Currie,J. and
Collura,K.
TITLE Rice Genomic Sequence
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 137327)
Wing,R.A., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Sasaki,C.,
Currie,J., Collura,K. and Thompson,S.
TITLE Direct Submision
JOURNAL Submitted (09-OCT-2002) Arizona Genomics Institute, University of
Arizona, 303 Forbes, Tucson, AZ 85721, USA
REFERENCE 3 (bases 1 to 137327)
Wing,R.A., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Sasaki,C.,
Currie,J., Collura,K. and Thompson,S.
TITLE Direct Submision
JOURNAL Submitted (06-NOV-2002) Arizona Genomics Institute, University of
Arizona, 303 Forbes, Tucson, AZ 85721, USA
REFERENCE 4 (bases 1 to 137327)
Wing,R.A., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Currie,J.,
Collura,K. and Thompson,S.
TITLE Direct Submision
JOURNAL Submitted (11-JAN-2003) Arizona Genomics Institute, University of
Arizona, 303 Forbes, Tucson, AZ 85721, USA
REFERENCE 5 (bases 1 to 137327)
Wing,R.A., Yu,Y., Soderlund,C., Chen,M., Kim,H., Rambo,T.,
Sasaki,C., Henry,D., Thompson,S., Simmons,J., Thurmond,S.K. and
Sun,S.
TITLE Direct Submision
JOURNAL Submitted (16-APR-2003) Clemson University Genomics Institute,
Clemson University, 100 Jordan Hall, Clemson, SC 29634, USA
COMMENT On Jan 11, 2003 this sequence version replaced gi:24635891.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality
>30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by more than
one plasmid subclone; and the assembly was confirmed by
restriction digest. There are transposons located at area
29427-32874 and at area 19945-33212. Region 70511-70536 is covered
by a single subclone. Region 18678-18873 is a double stranded
single subclone. Areas 69957-70761, 89836-90416, 98951-99009,
105316-105607, and 133052-133152 are covered by Monsanto only. The
nucleotide sequence of this BAC clone was generated by combing
Snygenta, Monsanto and Arizona Genomics Institute sequencing data.
Location/Qualifiers
1. 137327
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39947"
/chromosome="3"
/clone="OJ1041F02"
/note="(japonica cultivar-group)"
/complement(638, .9495)
/gene="OJ1041F02.1"
/note="unknown protein"
/complement(101n(638, .838,1141, .1276,3132, .3244,3326, .3421,
3513, .3727,3812, .3934,4024, .4054,4157, .4216,5088, .5243,

5322, .5411,5484, .5699,5808, .5942,6035, .6538,6622, .6837,
7548, .7593,7679, .7815,8196, .8328,9221, .9495))
/gene="OJ1041F02.1"
/codon_start=1
/product="unknown protein"
/protein_id="AAP06844.1"
/db_xref="GI:29893590"
/translation="MAQSLELLTIOFLMPDNDARROAEQIRRLADPPQVPAVYHNL
RTAKTPYRQAAVTLRKRTISHPKLPAPAKASIKAKLISITTDHSHLVRAANV
VSIKTAIVPAGRPFLPTIFQSQSFQEHREVALTLFSLETITGITQSHLDAN
OPTILKLCQDASSSRVRIALKAAGSFIEYVNDGDVVKIFRDFPVSILVNSQCLAN
GEBDVASTAFEPIDELIESPAFLGDSVRSIVQSFLEVSQNOELINIROAQIQLIS
LVKFAASFLKKGKLVIPILQVCCPLTATADGSDSLAARSAEVLINIDTMAINPRH
VPPPLTEASVSPRINRYKAEATVYSVSECCGCKKKECECLRVYELALDOE
OMNGAASPAFLGCPAEHQIPRLISYSEVLICINALADPDVEYKSYVALAACED
MGNLTPLTLPDMCKLVNSLOGSPRNLETCMSAIGVAAAEQFMFYAEVLVEMK
GPMVLTNDELCARARATEVGIYAMAAGRAMETILPFIETALISGVLDSELEY
THGFPSVNAEILIDSFAYLPHVVPVLAASCLNDGSAVDIDADSVNDSGVSSDD
DVNDEPRVNSIMLTITLADILTAIRAIIPRAADVLEKODILDTMNIYIKTMRDD
KEVVAOACTSLADIVRDCGFALIEPIYIRLADATLILRQSCCOOVSDGDDGDDID
HDEVLMADVSDLPAPFAKMSYDPPIFTKLPDSIMKRAKSPHPQDTMYATLAEV
AQMGKAPISAVYDKIMPLVLKELASSERNNAPFCVEMCKNGGAALATYGDILH
GLHRLFADEPDDAVRDNAAGAIAMINVOQSIPLNDILPVPVINAFAQVAVVSPN
ESDEVKTVAAVASHLISVGCQMPILSALPPAHNALASFANRR"
12372, .15503
/gene="OJ1041F02.2"
/note="unknown protein"
/join(12372, .12757,12858, .13025,13130, .13295,13716, .13826,
13982, .14130,14224, .14354,14459, .14558,15197, .15503)
/gene="OJ1041F02.2"
/codon_start=1
/product="unknown protein"
/protein_id="AAP06845.1"
/db_xref="GI:29893591"
/translation="MPQEGECRRHMLGVGEAAVGGSEEPFVPLRKRLSVDSKAS
TAPRICIWCDEGAEADITCDIYAAPLRSSCAKAMPAPAPFPMWTPPPRPQRDGE
BARPGEAIIKRGHSYSILMLQIGISYVSXSSALPQKLAASDPDRKRWTFPP
EGSKTTPPHSVDFRWKDYCPAFRHLKRLGVDAEYVLAICGDTLRELAIPKSGC
SCFPTODDRMIKTVKSEYKVLIRMRSLYEVHROVKSLTLIRFYCTIKACQCP
KIHRRFPLKSGSHGRTIDTKRKIDETTLTKDLDQYAFRLQRYEELMKOIQDCT
FLETQGLIDVSLGLGVHFRNDYVSXKIGISQHLAFPKSTGRKSEGGSSFCLECFE
SGCKDRDIDBRKPTIOLGIMPMQAEKSSKILDNPLNRRHLFTPPSGGSCDYL
PFGIIDLQDYDITKLEHAYKSFQVNDVYISAVDPKLSRRFQIFIRVFKEQ"
16734, .16878
/note="putative MITE, GalTjn/Gaigln-like"
19263, .19454
/note="(putative transposon, CACTA, En/Spm-like"
21819, .24660
/gene="OJ1041F02.3"
/note="(hypothetical protein"
/join(21819, .22007,22096, .22206,23772, .23996,24523, .24660)
/gene="OJ1041F02.3"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAP06846.1"
/db_xref="GI:29893592"
/translation="MATASDSSGGSGGEORRLISTPKESGRTIAPTRPDGTLRKAIRI
RAGYVPOEVAIYQSKAQMKRSGPDVPGYDPAIDAPKTKAKAKNRKRRKQAS
TTNDGKGLHIEDAGETDNPKNADVSTKQISGLAISLSLVATSTADATNKSSES
SAPDIDKKIRALKKKIRLAEAOVGDPENLKPEOLEKMKKIEGMKEELKLENNKSSPA
AS"
complement(26563, .30575)
/gene="OJ1041F02.4"
/note="unknown protein"
complement(101n(26563, .26772,27082, .27288,27430, .27537,
27881, .27940,28051, .28126,28327, .28484,28616, .28678,
28790, .28874,29560, .30575))
/gene="OJ1041F02.4"
/codon_start=1
/product="unknown protein"
/protein_id="AAP06847.1"
/db_xref="GI:29893593"
/translation="MGPPAAPSPSSSSSGSSRRRRRLRLRRNNAKHICYDASNFCAYP

KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS FULLTOP.
 SOURCE Danio rerio (zebrafish)
 ORGANISM Danio rerio
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
 REFERENCE 1 (bases 1 to 174806)
 AUTHORS McLaren, S.
 TITLE Direct Submission
 JOURNAL Submitted (28-SEP-2004) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Sep 25, 2004 this sequence version replaced gi:42821090.
 COMMENT ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: http://www.sanger.ac.uk
 Contact: zfish-help@sanger.ac.uk
 Project Information
 Center project name: zc125M22
 ----- Summary Statistics
 Assembly program: XGAP4; version 4.5
 Chemistry: Dye-terminator; 100% of reads
 Consensus quality: 172622 bases at least Q40
 Consensus quality: 173174 bases at least Q30
 Consensus quality: 173717 bases at least Q20
 Insert size: 174406; sum-of-contigs
 Insert size: 182163; 2.3% error; agarose-fp
 Quality coverage: 11.29x in Q20 bases; sum-of-contigs Quality coverage: 11.03x in Q20 bases; agarose-fp

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 5 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of 'N', but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 *
 1 74620: contig of 74620 bp in length
 * 74621 74720: gap of 100 bp
 * 74721 133368: contig of 58648 bp in length
 * 133369 133468: gap of 100 bp
 * 133469 150793: contig of 17325 bp in length
 * 150794 150893: gap of 100 bp
 * 150894 172651: contig of 21758 bp in length
 * 172652 172752: gap of 100 bp
 * 172752 174806: contig of 2055 bp in length.
 Location/Qualifiers
 source
 1. 174806
 /organism="Danio rerio"
 /mol_type="genomic DNA"
 /db_xref="taxon:7955"
 /clone="CH211-125M22"
 /clone_lib="CHORI-211"
 1. 74620
 /note="assembly_fragment:02018
 fragment_chain:1
 clone_end:586
 vector_side:left"
 74721..133368
 /note="assembly_fragment:00796
 fragment_chain:1"
 133469..150793
 /note="assembly_fragment:00037
 fragment_chain:1"
 150894..172651
 /note="assembly_fragment:00366
 fragment_chain:1"
 172752..174806
 /note="assembly_fragment:00028
 clone_end:17
 vector_side:right"

ORIGIN
 Query Match 22.2%; Score 34.8; DB 2; Length 174806;
 Best Local Similarity 52.8%; Pred. No. 2.7;
 Matches 75; Conservative 0; Mismatches 67; Indels 0; Gaps 0;
 Oy 15 TGGCATCTACATGCTGTTCACACCGTCGCTTGTTCATCGTCGAAGCTTGGCTATT 74
 Db 162638 TGGGTTGTTCATCTACTTACCCAGCGTGGGTTGTAAACACACCGCTTTAAAGT 162639
 Oy 75 CTGAACCAAGAGATACCTACTCCCAACAATTCATCTCATGATGCAACTTCATGCA 134
 Db 162638 AAGTAACATTCATCTAAAAAGCCCTAAATATTCTTTAAATCTGAATCTCAGCTTAA 162679
 Oy 135 ACACGCACATATGTTCTCGAA 156
 Db 162578 ACTCCACATTCATCTTAA 162557
 RESULT 12
 AC104134
 LOCUS AC104134 110127 bp DNA linear PRI 20-MAR-2002
 DEFINITION Homo sapiens BAC clone RP11-525L16 from 2, complete sequence.
 AC104134
 ACCESSION AC104134.4 GI:19551199
 VERSION AC104134.4 GI:19551199
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 110127)
 AUTHORS Sulston, J.E. and Waterston, R.
 TITLE Toward a complete human genome sequence
 JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
 MEDLINE 99063792
 PUBMED 9847074
 REFERENCE 2 (bases 1 to 110127)
 AUTHORS Tomlinson, C., Haekenson, W. and Doeber, A.
 TITLE The sequence of Homo sapiens BAC clone RP11-525L16
 JOURNAL Unpublished (2001)
 REFERENCE 3 (bases 1 to 110127)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (04-DEC-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 4 (bases 1 to 110127)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (15-FEB-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 5 (bases 1 to 110127)
 AUTHORS Waterston, R.
 TITLE Direct Submission
 JOURNAL Submitted (20-MAR-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA On Mar 20, 2002 this sequence version replaced gi:18677650.
 COMMENT ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc
 Contact: sapient@wustl.wustl.edu
 ----- Summary Statistics
 Center project name: H_NH0525L16

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:

all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCT-11 human BAC library was made from the blood of one male donor, as described by Osogawa, K., Moon, P.Y., Zhao, B., Frengsen, E., Tateno, M., Cataneese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.regen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-554H10, 2000 bp overlap; the clone sequenced to the right is RP11-450B3, 2000 bp overlap. Actual start of this clone is at base position 151956 of RP11-554H10.

FEATURES
source

```
1. 110127
   location/Qualifiers
   /organism="Homo sapiens"
   /mol_type="genomic DNA"
   /db_xref="taxon:9606"
   /chromosome="2"
   /map="2"
   /clone="RP11-525L16"
   /clone_1b="RPCT-11"
   701..998
   repeat_region
   /rpt_family="Alu"
   999..1297
   repeat_region
   /rpt_family="Alu"
   1410..1486
   repeat_region
   /rpt_family="L1"
   1564..1678
   repeat_region
   /rpt_family="Alu"
   1684..1722
   repeat_region
   /rpt_family="AT-rich"
   2338..2456
   repeat_region
   /rpt_family="MIR"
   4039..4332
   repeat_region
   /rpt_family="Alu"
   4412..4532
   repeat_region
   /rpt_family="Alu"
   4669..4846
   repeat_region
   /rpt_family="MIR"
   5878..6173
   repeat_region
   /rpt_family="Alu"
   6385..6584
   repeat_region
   /rpt_family="MIR"
   6929..7121
   repeat_region
   /rpt_family="MIR"
   7167..7312
   repeat_region
   /rpt_family="MIR"
   7953..8202
   repeat_region
   /rpt_family="Alu"
   8537..8577
   repeat_region
   /rpt_family="L2"
   8585..8756
   repeat_region
   /rpt_family="L1"
   8768..8888
   repeat_region
   /rpt_family="Alu"
   9005..9219
   repeat_region
   /rpt_family="L2"
```

```
repeat_region 9272..9838
   /rpt_family="L1"
repeat_region 9849..9976
   /rpt_family="Alu"
repeat_region 10124..10411
   /rpt_family="Alu"
repeat_region 10483..10759
   /rpt_family="L2"
repeat_region 11215..11264
   /rpt_family="(TA)n"
repeat_region 11362..11476
   /rpt_family="Alu"
repeat_region 11478..11776
   /rpt_family="Alu"
repeat_region 12314..12607
   /rpt_family="Alu"
repeat_region 12778..12943
   /rpt_family="MALR"
repeat_region 12954..12979
   /rpt_family="Alu"
repeat_region 14391..14655
   /rpt_family="Alu"
repeat_region 14670..14939
   /rpt_family="MALR"
repeat_region 15439..15449
   /rpt_family="L2"
repeat_region 15450..15730
   /rpt_family="Alu"
repeat_region 15731..15824
   /rpt_family="L2"
repeat_region 15837..15915
   /rpt_family="MALR"
repeat_region 16150..16214
   /rpt_family="MALR"
repeat_region 16224..16382
   /rpt_family="MER1_type"
repeat_region 16533..16747
   /rpt_family="MIR"
repeat_region 16889..17116
   /rpt_family="L2"
repeat_region 17271..17343
   /rpt_family="L1"
repeat_region 17384..17534
   /rpt_family="MIR"
repeat_region 17545..17843
   /rpt_family="Alu"
repeat_region 17995..18056
   /rpt_family="L2"
repeat_region 18057..18074
   /rpt_family="GA-rich"
repeat_region 18075..18379
   /rpt_family="Alu"
repeat_region 18380..18404
   /rpt_family="GA-rich"
repeat_region 18889..19067
   /rpt_family="MER1_type"
repeat_region 19296..19302
   /rpt_family="Alu"
repeat_region 19593..19709
   /rpt_family="Alu"
repeat_region 19890..20198
   /rpt_family="Alu"
repeat_region 20225..20698
   /rpt_family="L1"
repeat_region 20737..20769
   /rpt_family="AT-rich"
repeat_region 20887..21186
   /rpt_family="Alu"
misc_feature 21350..21369
   /note="match to EST BG571642 (NID:913579295)"
repeat_region 21357..21652
   /rpt_family="Alu"
repeat_region 21653..21957
```

```

repeat_region      22328..22636
                    /rptc_family="Alu"
Query Match      22.0%; Score 34.6; DB 9; Length 110127;
Best Local Similarity 54.3%; Pired. No. 3.1;
Matches 70; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

Qy      3 TGTGATCTCATGATGCGATCTACTATGCTGTTCACCGCTTCCTTGTTCATCTGTCGCA 62
Db      8763 TGGGATTACAGAGGTAGACGACGTGACCTGGCCACCTATCTTTCTCTCTCTCCCA 87692

Qy      63 GCGTTGCTTATTCTGAACCAAGAGATCTACTCTCCAAACAATTCATCTTACTACGCA 122
Db      87693 TCTCCCTCTTATCTTACCTCTTGTTCATCTACTCAGAACTCACAATTCCTCATACC 87752

Qy      123 ACTTCGATG 131
Db      87753 ACCTCCAG 87761

```

```

RESULT 13
AC136896      214946 bp      DNA      linear      PRI 21-MAR-2003
LOCUS      Homo sapiens chromosome 15, clone RP11-86001, complete sequence.
DEFINITION
AC136896
VERSION
AC136896.6      GI:29135656
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

```

REFERENCE
AUTHORS      Birren,B., Nusbaum,C. and Lander,E.
TITLE      Homo sapiens chromosome 15, clone RP11-86001
JOURNAL
REFERENCE
AUTHORS      2 (bases 1 to 214946)

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS

```

Birren,B., Nusbaum,C., Lander,E., Allen,A., Allen,N., Anderson,S.,
 Barmen,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B.,
 Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
 Cook,A., Cooke,P., Corum,B., Dearellano,K., Diaz,J.S., Dodge,S.,
 Fero,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S.,
 Gord,S., Graham,L., Grand-Pierre,N., Hafez,N., Hafez,N.,
 Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
 Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
 Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J.,
 Matthews,C., McCarthy,M., Meldrim,J., Menues,L., Mihova,T.,
 Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
 Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
 Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
 Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupack,R.,
 Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
 Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K.,
 Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
 Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (09-NOV-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 214946)

```

TITLE
JOURNAL
REFERENCE
AUTHORS

```

```

TITLE
JOURNAL
COMMENT

```

```

FEATURES
source
    1..214946
        location/Qualifiers
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
        /chromosome="15"
        /map="15"
        /clone="RP11-86001"
        /clone_lib="RP11 Human Male BAC"
        /complement(1..362)
        /rptc_family="HML1"
        /complement(823..1431)
        /rptc_family="L1MCS"
        913..922
        /note="<30 qual SNGL region"
        /complement(2941..3295)
        /rptc_family="MER31B"
        /complement(3307..3311)
        /note="<30 qual SNGL region"
        3438..3476
        /note="single clone coverage"
        /note="single clone coverage"
        3546..3586
        /rptc_family="GA-rich"
        3601..4006
        /rptc_family="MT1A1"
        /complement(4008..4166)
        /rptc_family="MIR"
        /complement(4580..4711)

```

```

repeat_region      /rpt_family="L3"
4731..4751
/rpt_family="AT_rich"
repeat_region      complement(4758..4993)
/rpt_family="MIR"
repeat_region      complement(5336..5464)
/rpt_family="L3"
repeat_region      5935..6412
/rpt_family="MER39"
repeat_region      complement(6466..7069)
/rpt_family="MLT1E"
repeat_region      complement(8633..8713)
/rpt_family="MIR"
repeat_region      8768..9244
/rpt_family="L1M2"
repeat_region      complement(9299..9498)
/rpt_family="MIR"
repeat_region      10252..10587
/rpt_family="MLT1A0"
repeat_region      12232..12619
/rpt_family="L1M5"
repeat_region      complement(13183..13354)
/rpt_family="MIR"
repeat_region      14074..14500
/rpt_family="L1M3"
repeat_region      15070..15201
/rpt_family="FLAM_C"
repeat_region      complement(16057..18621)
/rpt_family="L1M3"
repeat_region      complement(18678..19377)
/rpt_family="L1M1"
repeat_region      complement(19375..20142)
/rpt_family="L1M3"
repeat_region      22615..22919
/rpt_family="A1USg"
repeat_region      complement(25101..25206)
/rpt_family="HAL1"
repeat_region      complement(25691..25925)
/rpt_family="HAL1"
repeat_region      25996..26033
/rpt_family="(CAGA)n"
repeat_region      complement(26059..26582)
/rpt_family="MLT2B3"
repeat_region      26645..27455
/rpt_family="MER21B"
repeat_region      complement(27582..27833)
/rpt_family="HAL1"
repeat_region      complement(28185..28759)
/rpt_family="LTR68"
repeat_region      complement(29229..29603)
/rpt_family="THE1A"
repeat_region      complement(29771..30036)
/rpt_family="A1Ub"
repeat_region      30056..30083
/rpt_family="(TTTA)n"
repeat_region      complement(30090..30369)
/rpt_family="A1Udo"
repeat_region      31042..32585
/rpt_family="L1M4"
repeat_region      32586..32654
/rpt_family="(TCTA)n"
repeat_region      32655..33163
/rpt_family="L1M4"
repeat_region      33164..33473
/rpt_family="A1Uy"
repeat_region      33474..33683
/rpt_family="L1M4"
repeat_region      33686..33751
/rpt_family="MLT1D"
repeat_region      36941..37015
/rpt_family="MIR"
repeat_region      complement(38001..38449)
/rpt_family="L1MD1"

```

```

repeat_region      38453..38855
/rpt_family="MLT1D"
repeat_region      complement(38856..39656)

Query Match      21.9%; Score 34.4; DB 9; Length 214946;
Best Local Similarity 54.8%; Pred. No. 3.7;
Matches 68; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

Qy      5 TCATCTCAGTCGACTACTACATGCTTGTCAACCGTTCCTTGTTCATGCTCCAGC 64
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      35745 TCATCTCTATGGGACAGAAAAATCATTTGACAAAATTAGCATCTTCATGATGTAAAC 35804
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

Qy      65 CTTCCTATTTCGAACCAAGAGATACCTACTCCCAAAACATTCATCTACTATGCAAC 124
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      35805 ATCCAACTACTTAAGAAAGAAAGTGTCTTTTAAACCAATTAATGTAACCTATGAAAA 35864
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

Qy      125 TTCC 128
|||||
Db      35865 ATCC 35868

```

```

RESULT 14
CR847532/c      141810 bp      DNA      linear      HTG 03-OCT-2004
LOCUS
DEFINITION      Danio rerio clone DKEXP-86C9, *** SEQUENCING IN PROGRESS ***, 8
unordered pieces.
ACCESSION      CR847532
VERSION      CR847532.2 GI:53755907
KEYWORDS      HTG; HTGS PHASE1.
SOURCE      Danio rerio (zebrafish)
ORGANISM      Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 141810)
McLay, K.
Direct Submission
Submitted (02-OCT-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Oct 3, 2004 this sequence version replaced gi:53145936.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: zK086C9
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 13909 bases at least Q40
Consensus quality: 140415 bases at least Q30
Consensus quality: 140735 bases at least Q20
Insert size: 14110; sum-of-contigs
Insert size: 140600; 2.5% error; agarose-fp
Quality coverage: 8.52x in Q20 bases; sum-of-contigs Quality
coverage: 8.64x in Q20 bases; agarose-fp
-----

```

```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 19172: contig of 19172 bp in length
* 19173 19272: gap of 100 bp
* 19273 46769: contig of 27497 bp in length
* 46770 46869: gap of 100 bp
* 46870 52415: contig of 5546 bp in length
* 52416 52515: gap of 100 bp

```

* 52516 64328: contig of 11813 bp in length
* 64329 64428: gap of 100 bp
* 64429 96224: contig of 31796 bp in length
* 96225 96324: gap of 100 bp
* 96325 121956: contig of 25632 bp in length
* 121957 122056: gap of 100 bp
* 122057 131332: contig of 9276 bp in length
* 131333 131432: gap of 100 bp
* 131433 141810: contig of 10378 bp in length.

FEATURES

Source
1. 141810
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEYP-86C9"
/clone_1lb="DanioKeypliot"
1. 19172
/note="assembly fragment:00300
fragment chain:1"
19273. 46769
/note="assembly fragment:00469
fragment chain:1"
46870. 52415
/note="assembly fragment:00017
fragment chain:1"
52516. 64328
/note="assembly fragment:00167
fragment chain:1"
64429. 96224
/note="assembly fragment:00976
fragment chain:2"
96325. 121956
/note="assembly fragment:01302
fragment chain:2"
122057. 131332
/note="assembly fragment:00056.0"
131433. 141810
/note="assembly fragment:00703"

ORIGIN

Query Match 21.5%; Score 33.8; DB 2; Length 141810;
Best Local Similarity 52.5%; Pred. No. 5.8;

Matches 74; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 12 ACATGGCATCTACATGCTTGTTCACACCGTTCGCTTGTTCATCGCCAAAGCCTTGCCCT 71
DB 41846 ACAGTGAACCCCTTACATCTTGTGCTTCATGATCATATTGTACAGTGAAGAGCTTGGGT 41787
QY 72 ATTCTGAACCAAGAGATACCTATCTCCAAACAATCCATCTTACTCATGCAACTTCCATG 131
DB 41786 AAGTGGGAATATGAGGATCATCTGCTCTGTGAAAAATGAAATTAATACACCAAGAAACTT 41727
QY 132 CAACACGACATATGTTTCC 152
DB 41726 TCAAAAAAACTTATTAGTCC 41706

RESULT 15

BX640577 214178 bp DNA linear VRT 10-JUL-2004
LOCUS zebrafish DNA sequence from clone DKEY-117014 in linkage group 14,
DEFINITION complete sequence.
ACCESSION BX640577
VERSION BX640577.17 GI:50080105
KEYWORDS HTG.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 214178)
AUTHORS McGuire,S.
TITLE Direct Submission

JOURNAL

COMMENT

Submitted (10-JUL-2004) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Jul 8, 2004 this sequence version replaced gi:49659408.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: zfish-help@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

Em: EMBL, Sw.: SWISSPROT, Tr.: TREMBL, Wp.: WormPeP, Information on the WormPeP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats. Where this is found the longest good quality representation will be submitted.

Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhifeng Bao and Sean Bddy, submitted), and those beginning 'dr' were identified by Rick Waterman (Stephen Johnson lab, Washu). For further information see http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml DKEY-117014 is from a zebrafish BAC library

VECTOR: pindigobac-5.

FEATURES

Source
1. 214178
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-117014"
/clone_1lb="DanioKey"

ORIGIN

Query Match 21.5%; Score 33.8; DB 5; Length 214178;
Best Local Similarity 52.5%; Pred. No. 5.9;

Matches 74; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 12 ACATGGCATCTACATGCTTGTTCACACCGTTCGCTTGTTCATCGCCAAAGCCTTGCCCT 71
DB 58301 ACAGTGAACCCCTTACATCTTGTGCTTCATGATCATATTGTACAGTGAAGAGTGGGT 58360
QY 72 ATTCTGAACCAAGAGATACCTATCTCCAAACAATCCATCTTACTCATGCAACTTCCATG 131
DB 58361 AAGTGGGAATATGAGGATCATCTGCTCTGTGAAAAATGAAATTAATACACCAAGAAACTT 58420
QY 132 CAACACGACATATGTTTCC 152
DB 58421 TCAAAAAAACTTATTAGTCC 58441

Search completed: September 16, 2005, 03:01:19
Job time : 531.823 secs

XX Sequence 255 BP; 59 A; 86 C; 39 G; 71 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 157; DB 5; Length 255;
Best Local Similarity 100.0%; Pred. No. 1.3e-42;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGTGTATCTCATGAGCATGACTGATGCTGTTCACCGTTGCTGTTCATGCTCC 60
DB 15 CGTGTATCTCATGAGCATGACTGATGCTGTTCACCGTTGCTGTTCATGCTCC 74
QY 61 AAGCCTTGCTTATCTGAAACCAAGAGATACCTACTCCCAACATCTTACTCATG 120
DB 75 AAGCCTTGCTTATCTGAAACCAAGAGATACCTACTCCCAACATCTTACTCATG 134
QY 121 CAACCTTCATGCAAAACGACATATGTTCTCTGAAC 157
DB 135 CAACCTTCATGCAAAACGACATATGTTCTCTGAAC 171
RESULT 2
AAK07408 standard; DNA; 1394 BP.
AC AAK07408;
XX
XX 08-JUN-1999 (first entry)
XX
DE Zea mays Me45 male tissue-preferred regulatory region.
XX
XX Me45; male; tissue-preferred; regulatory region; plant cells;
KM plant tissue; differentiated; maize; hybrid seed; fertility; ss.
XX
OS Zea mays.
XX
XX WO9859061-A1.
XX
XX 30-DEC-1998.
XX
XX 19-JUN-1998; 98WO-US012895.
XX
XX 23-JUN-1997; 97US-00880499.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX PI Albertsen MC, Fox TW, Garnaat CW, Huffman GA, Kendall TL;
XX
XX MPI; 1999-105628/09.
XX
XX New nucleic acid encoding a Me45 male tissue-preferred regulatory region
PT - useful in mediating plant fertility, especially hybrid seed production.
XX
XX Claim 2; Page 22-23; 39pp; English.
XX
XX The sequence is that encoding an Me45 male tissue-preferred regulatory
CC region. It may be used in the construction of a vector for a method of
CC producing exogenous genes in a male tissue- preferred manner, which is
CC useful in restoring or conferring fertility, such as in hybrid seed
CC production. In conferring fertility, a monocot/dicot plant is transformed
CC with the exogenous nucleotide sequence (a male sterility gene, preferably
CC Me45), which encodes a product selected from auxins, rolB and diptheria
CC toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
CC and infertile plants
XX
SQ Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;
Query Match 100.0%; Score 157; DB 2; Length 1394;
Best Local Similarity 100.0%; Pred. No. 2.1e-42;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGTGTATCTCATGAGCATGACTGATGCTGTTCACCGTTGCTGTTCATGCTCC 60
DB 1155 CGTGTATCTCATGAGCATGACTGATGCTGTTCACCGTTGCTGTTCATGCTCC 1214

QY 61 AAGCCTTGCTTATCTGAAACCAAGAGATACCTACTCCCAACATCTTACTCATG 120
DB 1215 AAGCCTTGCTTATCTGAAACCAAGAGATACCTACTCCCAACATCTTACTCATG 1274
QY 121 CAACCTTCATGCAAAACGACATATGTTCTCTGAAC 157
DB 1275 CAACCTTCATGCAAAACGACATATGTTCTCTGAAC 1311
RESULT 3
AAK07409 standard; DNA; 1394 BP.
ID AAK07409
XX
XX AAK07409;
XX
XX 08-JUN-1999 (first entry)
XX
XX
DE Zea mays Me45 male tissue-preferred regulatory region.
XX
XX Me45; male; tissue-preferred; regulatory region; plant cells;
KM plant tissue; differentiated; hybrid seed; fertility; ss.
XX
OS Zea mays.
XX
XX WO9859061-A1.
XX
XX 30-DEC-1998.
XX
XX 19-JUN-1998; 98WO-US012895.
XX
XX 23-JUN-1997; 97US-00880499.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX PI Albertsen MC, Fox TW, Garnaat CW, Huffman GA, Kendall TL;
XX
XX MPI; 1999-105628/09.
XX
XX New nucleic acid encoding a Me45 male tissue-preferred regulatory region
PT - useful in mediating plant fertility, especially hybrid seed production.
XX
XX Claim 3; Page 23-24; 39pp; English.
XX
XX The sequence is that encoding an Me45 male tissue-preferred regulatory
CC region. It may be used in the construction of a vector for a method of
CC producing exogenous genes in a male tissue- preferred manner, which is
CC useful in restoring or conferring fertility, such as in hybrid seed
CC production. In conferring fertility, a monocot/dicot plant is transformed
CC with the exogenous nucleotide sequence (a male sterility gene, preferably
CC Me45), which encodes a product selected from auxins, rolB and diptheria
CC toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
CC and infertile plants
XX
SQ Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;
Query Match 100.0%; Score 157; DB 2; Length 1394;
Best Local Similarity 100.0%; Pred. No. 2.1e-42;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGTGTATCTCATGAGCATGACTGATGCTGTTCACCGTTGCTGTTCATGCTCC 60
DB 1155 CGTGTATCTCATGAGCATGACTGATGCTGTTCACCGTTGCTGTTCATGCTCC 1214
QY 61 AAGCCTTGCTTATCTGAAACCAAGAGATACCTACTCCCAACATCTTACTCATG 120
DB 1215 AAGCCTTGCTTATCTGAAACCAAGAGATACCTACTCCCAACATCTTACTCATG 1274
QY 121 CAACCTTCATGCAAAACGACATATGTTCTCTGAAC 157
DB 1275 CAACCTTCATGCAAAACGACATATGTTCTCTGAAC 1311

```
RESULT 4
AAH76332
ID AAH76332 standard; DNA; 1394 BP.
XX
XX
AC AAH76332;
XX
XX
DT 29-OCT-2001 (first entry)
XX
XX
DE 2. mayb Ms45 male tissue-preferred regulatory region encoding DNA.
XX
XX
KM Ms45; male tissue; regulatory region; transcription; male fertility;
XX
XX
KW hybrid seed; de.
XX
OS Zea mayb.
XX
PN WO200160997-A2.
XX
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
DR WPI; 2001-514772/56.
XX
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the Ms45 gene useful for
PT mediating fertility in a male plant.
XX
XX
PS Claim 4; Page 46; 50pp; English.
XX
XX
CC The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the Ms45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a nucleic acid sequence encoding an Ms45 male
CC tissue preferred regulatory region from Z. mayb
XX
XX
SQ Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;
Query Match 100.0%; Score 157; DB 5; Length 1394;
Best Local Similarity 100.0%; Pred. No. 2.1e-42;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGTGCATCTCACATGCGATACATGCTTGTTCACCGCTGCTGTTCCATGCTCC 60
DB 1155 CGTGCATCTCACATGCGATACATGCTTGTTCACCGCTGCTGTTCCATGCTCC 1214
QY 61 AAGCCTTGCTTATTTGGAACCAAGAGATACCTACTCCCAACAAATCCATCTTACTCATG 120
DB 1215 AAGCCTTGCTTATTTGGAACCAAGAGATACCTACTCCCAACAAATCCATCTTACTCATG 1274
QY 121 CAACCTTCATGCAAAACGCGACATATGTTTCTGTAAC 157
DB 1275 CAACCTTCATGCAAAACGCGACATATGTTTCTGTAAC 1311
RESULT 5
AAH76333
ID AAH76333 standard; DNA; 1394 BP.
XX
XX
AC AAH76333;
XX
XX
DT 29-OCT-2001 (first entry)
XX
XX
DE 2. mayb Ms45 male tissue-preferred regulatory region encoding DNA.
```

```
XX
XX
KM Ms45; male tissue; regulatory region; transcription; male fertility;
XX
XX
KW hybrid seed; de.
XX
XX
OS Zea mayb.
XX
PN WO200160997-A2.
XX
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
DR WPI; 2001-514772/56.
XX
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the Ms45 gene useful for
PT mediating fertility in a male plant.
XX
XX
PS Claim 4; Page 47; 50pp; English.
XX
XX
CC The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the Ms45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a nucleic acid sequence encoding an Ms45 male
CC tissue preferred regulatory region from Z. mayb
XX
XX
SQ Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;
Query Match 100.0%; Score 157; DB 5; Length 1394;
Best Local Similarity 100.0%; Pred. No. 2.1e-42;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGTGCATCTCACATGCGATACATGCTTGTTCACCGCTGCTGTTCCATGCTCC 60
DB 1155 CGTGCATCTCACATGCGATACATGCTTGTTCACCGCTGCTGTTCCATGCTCC 1214
QY 61 AAGCCTTGCTTATTTGGAACCAAGAGATACCTACTCCCAACAAATCCATCTTACTCATG 120
DB 1215 AAGCCTTGCTTATTTGGAACCAAGAGATACCTACTCCCAACAAATCCATCTTACTCATG 1274
QY 121 CAACCTTCATGCAAAACGCGACATATGTTTCTGTAAC 157
DB 1275 CAACCTTCATGCAAAACGCGACATATGTTTCTGTAAC 1311
RESULT 6
AAH76334
ID AAH76334 standard; DNA; 158 BP.
XX
XX
AC AAH76334;
XX
XX
DT 29-OCT-2001 (first entry)
XX
XX
DE 2. mayb Ms45 male tissue-preferred regulatory region fragment.
XX
XX
KM Ms45; male tissue; regulatory region; transcription; male fertility;
XX
XX
KW hybrid seed; de.
XX
OS Zea mayb.
XX
PN WO200160997-A2.
XX
PD 23-AUG-2001.
```


CC present sequence represents a DNA fragment upstream of the TATA box of a
CC 2. mays M45 male-tissue preferred regulatory region nucleotide sequence
XX
SQ Sequence 40 BP, 14 A, 14 C, 3 G, 9 T, 0 U, 0 Other;
Query Match 25.5%; Score 40; DB 5; Length 40;
Best Local Similarity 100.0%; Pred. No. 0.0015;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 85 AGGATACCTACTCCCAACATCCATCTTACTCATGCAAC 124
DB 1 AGGATACCTACTCCCAACATCCATCTTACTCATGCAAC 40
RESULT 9
ADK52131
ID ADK52131 standard; cDNA; 3267 BP.
XX
AC ADK52131;
XX
DT 20-MAY-2004 (first entry)
XX
DE Mouse atopic-dermatitis/psoriasis-associated EST #4.
XX
KM Mouse; 68; EST; atopic dermatitis; psoriasis; dermatological;
XX anti-inflammatory; antipruritic; rash; expressed sequence tag.
XX
OS Mus musculus.
XX
PN W02004016785-A1.
XX
PD 26-FEB-2004.
XX
PF 06-AUG-2003; 2003MO-JF009999.
XX
PR 06-AUG-2002; 2002JP-00229319.
XX 14-MAY-2003; 2003JP-00136544.
XX
PA (GENO-) GENOX RES INC.
XX (U9JU-) UNIV JUNTENDO.
XX
PI Itoh M, Ogawa K, Shinagawa A, Sudo H, Ogawa H, Ra C;
XX Mitsuiishi K;
XX WPI; 2004-214514/20.
XX
DR
XX
PT Detecting atopic dermatitis or psoriasis comprises assaying levels of
XX expression of an indicator gene at a rash site and non-rash site of a
XX person with atopic dermatitis or psoriasis.
XX
PS Claim 20; SEQ ID NO 164; 484bp; Japanese.
XX
XX The invention relates to detecting atopic dermatitis or psoriasis
XX comprising assaying the levels of expression of an indicator gene at a
XX rash site and non-rash site of a person with atopic dermatitis or
XX psoriasis, comparing these levels with those of a healthy person, and
XX determining that if the levels of indicators are higher or lower, then
XX this indicates the disease. Also included are a reagent for detecting
XX atopic dermatitis or psoriasis, a kit for screening for treatments, a
XX transgenic non human vertebrate animal models for the diseases, an agent
XX for inducing the diseases in mice and a DNA chip for assaying for the
XX indicator genes. The method is used for treatment, detection and animal
XX models for research of atopic dermatitis and psoriasis. The present
XX sequence is a Mouse atopic-dermatitis/ psoriasis-associated EST
XX (expressed sequence tag).
XX
SQ Sequence 3267 BP, 1029 A, 658 C, 590 G, 990 T, 0 U, 0 Other;
Query Match 20.6%; Score 32.4; DB 12; Length 3267;
Best Local Similarity 54.1%; Pred. No. 2.2;
Matches 66; Conservative 0; Mismatches 56; Indels 0; Gaps 0;
QY 27 TGGCTGTCAACCGTTCGTCTGTCATGTCACAGCCTGCTATTGTGAAACCAAGG 86

DB 2774 TTTTTCCTACCTCCCTCTGATTTCTTTGACCTCAGCTTTTATTTGGCCCTTTAAGTT 2833
QY 87 GATACCTACTCCCAACAAATCCATCTTACTCATGCAACTTCACGCAACACGACATAT 146
DB 2834 GTTCTAGATTCCCAAGAAAGCCATTTTGGACCTACAGTCAAGACTATTAACACCTGAAAT 2893
QY 147 GT 148
DB 2894 GT 2895
RESULT 10
ACA48402/C
ID ACA48402 standard; DNA; 2352 BP.
XX
AC ACA48402;
XX
DT 19-JUN-2003 (first entry)
XX
DE Prokaryotic essential gene #30059.
XX
KM Antisense; ds; prokaryotic essential gene; cell proliferation;
XX drug design; gene.
XX
OS Streptococcus mutans.
XX
PN W0200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002MO-US009107.
XX
XX 21-MAR-2001; 2001US-00815242.
XX 06-SEP-2001; 2001US-00948993.
XX 25-OCT-2001; 2001US-0342923P.
XX 08-FEB-2002; 2002US-00072851.
XX 06-MAR-2002; 2002US-0362899P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zykkind JW;
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
XX P-PSDB; ABU44532.
XX
DR
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation or
XX isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 14; SEQ ID NO 36272; 1766bp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX cell proliferation; (7) identifying a compound that influences the activity of
XX the gene product or that has an activity against a biological pathway;
XX CC required for proliferation, or that inhibits cellular proliferation; (8)
XX identifying a gene required for cellular proliferation or the biological
XX pathway in which a proliferation-regulated gene or its gene product lies
XX or a gene on which the test compound that inhibits proliferation of an
XX organism acts; (9) manufacturing an antibody; (10) profiling a
XX compound's activity; (11) a culture comprising strains in which the gene
XX product is overexpressed or underexpressed; (12) determining the extent
XX to which each of the strains is present in a culture or collection of

CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2352 BP; 762 A; 401 C; 500 G; 689 T; 0 U; 0 Other;
XX
Query Match 20.1%; Score 31.6; DB 8; Length 2352;
Best Local Similarity 58.5%; Pred. No. 3.8;
Matches 55; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
XX
QY 62 AGCCTTGCCCTATTCGACCAAGAGATACCTACTCCCAACCAATCCTTCTCATGC 121
DB 234 ATCCTGACTGTGCCCAAGCCAAAGAGACATACCTCTGAAAATCTGTGCTACATCTGC 175
QY 122 AACTTCATGCAAAACGACGACATATGTTCTCGA 155
DB 174 AATTTCAGCAAAAGAACGTTCAATTTTCTCTGA 141
XX
RESULT 11
ABN80329/C
ID ABN80329 standard; DNA; 15734 BP.
XX
AC ABN80329;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human chemically modified disease associated gene SEQ ID NO 346.
XX
DE Human; development; homeobox gene; HOX; diabetes; cancer; apoptosis;
KM heart disease; epilepsy; histone deacetylation; muscular dystrophy;
KM dwarfism; single nucleotide polymorphism; SNP; cytosine methylation;
KM antidiabetic; cytostatic; anticonvulsant; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN MO200200927-A2.
XX
PD 03-JAN-2002.
XX
PF 02-JUL-2001; 2001WO-EP007536.
XX
PR 30-JUN-2000; 2000DE-01032529.
PR 01-SEP-2000; 2000DE-01043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI, 2002-130908/17.
XX
PT Novel nucleic acid useful for diagnosis and therapy of diseases
PT associated with development genes such as diabetes, comprises a sequence
PT of a segment of chemically pretreated DNA of genes associated with
PT development.
XX
PS Claim 1; SEQ ID NO 346; 27pp; English.
XX
XX The invention relates to a nucleic acid (I) comprising a sequence at
CC least 18 bases in length of a segment of chemically pretreated DNA (II)
CC of genes associated with development selected from 87 genes listed in the
CC specification such as ACCPN, ADFN, or AFDI and comprising one of 350
CC sequences (ABN79984-ABN80333) or their complements. The invention is
CC useful for the diagnosis or therapy of diseases associated with

CC development genes, in particular disease related to homeobox containing
CC genes (HOX), like diabetes, cancer, apoptosis related diseases, syndromes
CC associated with congenital heart disease, epilepsy, diseases related to
CC histone deacetylation, Currarino syndrome, diseases related with the
CC development of the brain and limb girdle muscular dystrophy and dwarfism.
CC Oligomers specific to each of the genes are useful for detecting the
CC methylation state of all CpG dinucleotides within the 350 sequences or
CC (II) and their complementary sequences, as primer oligonucleotides for
CC the amplification of the 350 sequences, (II) and/or their complements and
CC as oligomer probes for detecting the cytosine methylation state and/or
CC single nucleotide polymorphisms (SNPs). Note: The sequence data for this
CC patent did not form part of the printed specification but is based on
CC sequence information supplied to Derwent by the European Patent Office
XX
SQ Sequence 15734 BP; 3630 A; 744 C; 4234 G; 7126 T; 0 U; 0 Other;
XX
Query Match 20.0%; Score 31.4; DB 6; Length 15734;
Best Local Similarity 59.6%; Pred. No. 8;
Matches 53; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
XX
QY 52 CCATGTCACACCTTGCTTATTCGACCAAGAGATACCTATCCCAATCCATCCATC 111
DB 8802 CTTCTTTCTCCCTATTCGAAATCTTAACATCAAAACCTTAATAATTAACACCAAC 8743
QY 112 TTACTCATGCAACTTCATGCAACACCGC 140
DB 8742 TACCTAACCCATCTTCCAAATTAACACCC 8714
XX
RESULT 12
AAF22279
ID AAF22279 standard; DNA; 64415 BP.
XX
AC AAF22279;
XX
DT 20-MAR-2001 (first entry)
XX
DE BAC containing repeats from centromeres 1-4 #2.
XX
DE Centromere; mitochondria; vector; ds.
XX
KM Arabidopsis thaliana.
XX
OS Arabidopsis thaliana.
OS Synthetic.
XX
PN WO200055325-A2.
XX
PD 21-SEP-2000.
XX
PF 17-MAR-2000; 2000WO-US007392.
XX
PR 18-MAR-1999; 99US-0125219P.
PR 01-APR-1999; 99US-0127409P.
PR 18-MAY-1999; 99US-0134770P.
PR 13-SEP-1999; 99US-0153584P.
PR 17-SEP-1999; 99US-0154603P.
PR 16-DEC-1999; 99US-0172493P.
XX
PA (UYCH-) UNIV CHICAGO.
XX
PI Preuss D, Copenhaver G, Keith K;
XX
DR WPI, 2000-587529/55.
XX
PT Recombinant DNA construct comprising a plant centromere, useful for
PT producing stably inherited mitochondria which can serve as vectors for the
PT construction of transgenic plant and animal cells.
XX
PS Claim 102; Page 321-335; 1449pp; English.
XX
XX The present invention relates to a recombinant DNA construct of a plant
CC (Arabidopsis thaliana) centromere. The constructs are useful for
CC producing stably inherited mitochondria which can serve as vectors for the
CC construction of transgenic plant and animal cells expressing selected
CC proteins such as hormones, enzymes, interleukins, clotting factors,

CC cytokines, antibodies, and growth factors
XX
SQ Sequence 64415 BP; 18698 A; 13554 C; 13083 G; 18980 T; 0 U; 100 Other;
Query Match 19.5%; Score 30.6; DB 3; Length 64415;
Best Local Similarity 58.1%; Pred. No. 23;
Matches 57; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
QY 65 CTTCCTATTCTGAAACCAAGAGATACCTACTCCGAACATCCATCTTACTATGCAAC 124
DB 7117 CTTCCTATTCTGAAATATATATATATTTGCTCTTAACACTAAACCTTAACCTACAC 7176
QY 125 TTCCATGCAACACGCAATATGTTTCTGAC 157
DB 7177 TTAAATCCAAACCTTAATCTAATTCCTTAAC 7209
RESULT 13
ADA02717
ID ADA02717 standard; DNA; 79467 BP.
XX
XX ADA02717;
XX
XX
DT 06-NOV-2003 (first entry)
XX
DE Mouse Nfatc1 carcinoma associated gene, SEQ ID NO:1235.
XX
XX Mouse; murine; carcinoma associated; oncogene; carcinoma; cancer; breast;
KM prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;
KM gene; de.
XX
XX Mus sp.
XX
XX WO2003057146-A2.
XX
XX 17-JUL-2003.
XX
XX 26-DEC-2002; 2002WO-US041414.
XX
XX 26-DEC-2001; 2001US-00035832.
XX
XX (SAGR-) SAGRES DISCOVERY.
XX
XX Morris DW;
XX
XX MPI; 2003-587068/55.
XX
XX
PT New recombinant nucleic acid encoding carcinoma associated protein,
PT useful for preparing compositions for treating carcinomas.
XX
XX
PS Claim 1; SEQ ID NO 1235; 245bp; English.
XX
XX The invention relates to recombinant carcinoma associated (CA) nucleic
XX acid sequences from mouse and human (ADA01482-ADA03094) and to
XX recombinant carcinoma associated proteins (CAP) encoded by them. The
XX invention also encompasses expression vectors and host cells comprising a
XX CA nucleic acid, a polypeptide (especially an antibody) that specifically
XX binds to the protein, and a biochip comprising CA nucleic acid or
XX fragments thereof. The sequences of the invention were identified using
XX oncogenic retroviruses, which insert into the genome of the host organism
XX at random. Many of these do not carry transduced host oncogenes or
XX pathogenic trans-acting viral genes, meaning that cancer incidence is a
XX direct consequence of the effects of proviral integration into host
XX protooncogenes. The CA nucleic acid sequences can be used to diagnose
XX carcinoma (especially breast cancer, prostate cancer, lymphoma or
XX leukemia) or a propensity to carcinoma by determination of the sequence
XX of a CA gene, or by determination of CA gene expression in particular
XX tissues. CA nucleic acids, proteins and antibodies are also useful as
XX therapeutic agents and in screening and evaluating drug candidates. The
XX present sequence represents a specifically claimed murine CA nucleic acid
XX sequence of the invention. Note: The complete sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX
SQ Sequence 79467 BP; 18256 A; 19684 C; 20062 G; 20204 T; 0 U; 1261 Other;
Query Match 19.5%; Score 30.6; DB 9; Length 79467;
Best Local Similarity 56.4%; Pred. No. 25;
Matches 57; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
QY 51 TCCATGCTCCAGCCTTGCTTATTTGAAACCAAGAGATACCTACTCCCAACATTCAT 110
DB 53779 TCCATTAATGATCTTGCTTCTCTGGAACAGCTGATCTTAATGTCACACACACACA 53838
QY 111 CTACTCATGCACTTCCATGCAACACGCAATATGTTTC 151
DB 53839 CACACACACACACACACACACACACACAGCTTGCTGCTC 53879
RESULT 14
ADB72455
ID ADB72455 standard; DNA; 79467 BP.
XX
XX ADB72455;
XX
XX
DT 04-DEC-2003 (first entry)
XX
DE Mouse Nfatc1 gene.
XX
XX mouse; ds; cytostatic; gene therapy; vaccine; carcinoma; lymphomas;
KM cancer; neoplasm; adenocarcinoma; sarcoma; gene.
XX
XX
XX Mus sp.
XX
XX WO2003008583-A2.
XX
XX 30-JAN-2003.
XX
XX 26-DEC-2001; 2001WO-US051291.
XX
XX 02-MAR-2001; 2001US-00798586.
XX
XX 23-OCT-2001; 2001US-0004113.
XX
XX 08-NOV-2001; 2001US-00052482.
XX
XX 30-NOV-2001; 2001US-00997722.
XX
XX 20-DEC-2001; 2001US-00034650.
XX
XX (SAGR-) SAGRES DISCOVERY.
XX
XX
XX Morris DW, Engelhard EK;
XX
XX MPI; 2003-239337/23.
XX
XX
PT New recombinant nucleic acid, useful for treating carcinomas, lymphomas,
PT cancers, neoplasm, adenocarcinoma, and sarcomas.
XX
XX
PS Claim 1; SEQ ID NO 283; 2304bp; English.
XX
XX The invention relates to a novel recombinant nucleic acid comprising a
XX nucleotide sequence selected from any of the 660 sequences fully defined
XX in the specification. A polynucleotide of the invention has cytostatic
XX activity, and may have a use in gene therapy, or in a vaccine. The
XX recombinant nucleic acids and polypeptides are useful for treating
XX carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and
XX sarcomas. The present sequence represents a mouse gene of the invention.
XX
XX
SQ Sequence 79467 BP; 18256 A; 19684 C; 20062 G; 20204 T; 0 U; 1261 Other;
Query Match 19.5%; Score 30.6; DB 10; Length 79467;
Best Local Similarity 56.4%; Pred. No. 25;
Matches 57; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
QY 51 TCCATGCTCCAGCCTTGCTTATTTGAAACCAAGAGATACCTACTCCCAACATTCAT 110
DB 53779 TCCATTAATGATCTTGCTTCTCTGGAACAGCTGATCTTAATGTCACACACACACA 53838

QY 111 CTTACTCATGCACTTCCATGCAACGACGACATATGTTTC 151
 Db 53839 CACACACACACACACACACACACACACACAGCTTGCGCTC 53879

RESULT 15

ADE95965
 ID ADE95965 standard; DNA; 79467 BP.

XX ADE95965;

XX 12-FEB-2004 (first entry)

XX Mouse Nfatc1 gene genomic DNA sequence.

XX
 KM cancer diagnosis; cancer treatment; carcinoma; cytostatic; gene therapy;
 KM lymphoma; breast cancer; prostate cancer; leukemia; ds; mouse; murine;
 KM Nfatc1.

XX Mus sp.

XX WO2003039484-A2.

XX 15-MAY-2003.

XX 08-NOV-2002; 2002WO-US036071.

XX 08-NOV-2001; 2001US-00052482.

XX (SAGR-) SAGRES DISCOVERY.

XX Morris DW, Engelhard EK;

XX MPI; 2003-441462/41.

XX New carcinoma associated nucleic acids and proteins, useful for screening
 PT drug candidates, or for diagnosing and treating carcinomas, e.g.
 PT lymphoma, breast cancer, prostate cancer or leukemia.

PS Claim 1; SEQ ID NO 223; 793pp; English.

CC This invention relates to novel recombinant nucleic acids for use in
 CC diagnosis and treatment of cancer, especially carcinomas, as well as the
 CC use of compositions in screening methods. The compositions of the
 CC invention may have cytostatic activity whilst the disclosed sequences may
 CC be useful for gene therapy. The carcinoma associated nucleic acids and
 CC proteins are useful for diagnosing and treating carcinomas, for example
 CC lymphoma, breast cancer, prostate cancer or leukemia, or for screening
 CC drug candidates or bioactive agents capable of binding to, or modulating
 CC the activity of, a carcinoma associated protein. The present sequence is
 CC the genomic DNA sequence of the mouse Nfatc1 gene which is a carcinoma
 CC associated gene of the invention.

XX Sequence 79467 BP; 18256 A; 19684 C; 20062 G; 20204 T; 0 U; 1261 Other;

Query Match 19.5%; Score 30.6; DB 10; Length 79467;

Best Local Similarity 56.4%; Pred. No. 25;

Matches 57; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 51 TCCATGCTCCAGCCTTGCTTATTTGACCAAGAGATACCTACTCCCAACATTCAT 110

Db 53779 TCCACTAAGTGATCTTGCTTCTGTGAAACAAGCTGAGCTTAGTCACACACACACA 53838

QY 111 CTTACTCATGCACTTCCATGCAACGACGACATATGTTTC 151

Db 53839 CACACACACACACACACACACACACACACAGCTTGCGCTC 53879

Search completed: September 15, 2005, 21:33:10
 Job time : 81.0101 secs

QY 95 CTCGCAACATCATCTTACTCA 118
DB 880 ATCCCTAGATGCTTTTATATA 857

RESULT 5

US-09-557-884-1/c
Sequence 1, Application US/09557884
Patent No. 6506581
GENERAL INFORMATION:
APPLICANT: Fleischmann et al.
TITLE OF INVENTION: The Nucleotide sequence of the Haemophilus influenzae Rd Genome, Fragments thereof, and Uses Thereof
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS v6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/557,884
FILING DATE: 25-Apr-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/476,102
FILING DATE: JUN-5-1995
ATTORNEY/AGENT INFORMATION:
NAME: Michelle S. Marks
REGISTRATION NUMBER: 41,971
REFERENCE/DOCKET NUMBER: PB186P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1830121 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1

Query Match 18.6%; Score 29.2; DB 4; Length 1830121;
Best Local Similarity 51.5%; Pred. No. 66;
Matches 67; Conservative 0; Mismatches 63; Indels 0; Gaps 0;
QY 12 ACATGCGATCACTACGCTTGTTCACCGTTCCTGCTTGTTCATCGTCCAAAGCTTGCC 71
DB 677275 ACTTTAAATTCAGCGCTGCTGCTGCTTATTCATTTGCAATTAAGCAAGATAGCC 677216
QY 72 ATTCTGAACCAAGAGATCACTACTCCCAACAATCATCTTACTCATGCACTTCATG 131
DB 677215 ATGCTGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 677156
QY 132 CAACACGCA 141
DB 677155 CAACCAACCA 677146

RESULT 6
US-09-643-990A-1/c
Sequence 1, Application US/09643990A
Patent No. 6528289
GENERAL INFORMATION:
APPLICANT: Robert D. Fleischmann

Mark D. Adams
Owen White
Hamilton O. Smith
J. Craig Venter

TITLE OF INVENTION: The Nucleotide sequence of the Haemophilus influenzae Rd Genome, Fragments thereof, and Uses Thereof

NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville,
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS v6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/643,990A
FILING DATE: 23-Aug-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/487,429
FILING DATE: 1995-06-07
APPLICATION NUMBER: 08/426,787
FILING DATE: 1995-04-21
ATTORNEY/AGENT INFORMATION:
NAME: Kenley K. Hoover
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: PB186P1C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-610-5790
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1830121 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1

Query Match 18.6%; Score 29.2; DB 4; Length 1830121;
Best Local Similarity 51.5%; Pred. No. 66;
Matches 67; Conservative 0; Mismatches 63; Indels 0; Gaps 0;
QY 12 ACATGCGATCACTACGCTTGTTCACCGTTCCTGCTTGTTCATCGTCCAAAGCTTGCC 71
DB 677275 ACTTTAAATTCAGCGCTGCTGCTGCTTATTCATTTGCAATTAAGCAAGATAGCC 677216
QY 72 ATTCTGAACCAAGAGATCACTACTCCCAACAATCATCTTACTCATGCACTTCATG 131
DB 677215 ATGCTGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 677156
QY 132 CAACACGCA 141
DB 677155 CAACCAACCA 677146

RESULT 7
US-09-949-016-14712/c
Sequence 14712, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14

```
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 14712
/ LENGTH: 64190
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-14712
```

```
Query Match      18.3%; Score 28.8; DB 4; Length 64190;
Best Local Similarity 62.5%; Pred. No. 24;
Matches 45; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
```

```
QY 55 TCGTCCAGCCTTCTATTTGACCAAGAGATACCTACCCAAACATCCATCTTA 114
DB 355 TCATCCAAATAGTTTCCACCTACTATCATCATGGAACGTAGTACTATCATCAAGCCATATT 296
QY 115 CTCATGCAACTT 126
DB 295 CTCATGTAAACAT 284
```

```
RESULT 8
US-09-949-016-14713/c
/ Sequence 14713, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 14713
/ LENGTH: 64190
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-14713
```

```
Query Match      18.3%; Score 28.8; DB 4; Length 64190;
Best Local Similarity 62.5%; Pred. No. 24;
Matches 45; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
```

```
QY 55 TCGTCCAGCCTTCTATTTGACCAAGAGATACCTACCCAAACATCCATCTTA 114
DB 355 TCATCCAAATAGTTTCCACCTACTATCATCATGGAACGTAGTACTATCATCAAGCCATATT 296
QY 115 CTCATGCAACTT 126
DB 295 CTCATGTAAACAT 284
```

```
RESULT 9
US-09-949-016-15502
/ Sequence 15502, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
```

```
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 15502
/ LENGTH: 11022
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-15502
```

```
Query Match      17.8%; Score 28; DB 4; Length 11022;
Best Local Similarity 63.2%; Pred. No. 22;
Matches 43; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
```

```
QY 15 TGGCATCTACATGCTTGTTGTCACCGTTCCTTGTTCATGTCACAGCTTGCTATT 74
DB 4432 TGGCTTATTTCATGGGTGTTTATGGGTTAATGTGTAGCATCATTTAGCCTTGTCAGG 4491
QY 75 CTGAACCA 82
DB 4492 CTGAATCA 4499
```

```
RESULT 10
US-09-949-016-17447
/ Sequence 17447, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 17447
/ LENGTH: 41815
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-17447
```

```
Query Match      17.8%; Score 28; DB 4; Length 41815;
Best Local Similarity 58.3%; Pred. No. 38;
Matches 49; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
```

```
QY 62 AGCCTTGCTATTTGGAACCAAGAGATACCTACTCCAAACATTCATTACTGATGC 121
DB 33058 AGACATGAGATTCACACACATCCGCATAGCTGCTCAACGATCCCTGGTCATCATTC 33117
QY 122 AACTCCATGCAAAACGACACATA 145
DB 33118 ATGTCAATAGACATGATGTA 33141
```

```
RESULT 11
US-09-949-016-191201/c
/ Sequence 191201, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
```


US-09-134-001C-1413/C
; Sequence 1413, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1413
; LENGTH: 1428
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1413

Query Match 17.5%; Score 27.4; DB 3; Length 1428;
Best Local Similarity 55.9%; Pred. No. 15;
Matches 52; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 14 ATGCATACCTACATGCTGTGTTCAACCGTTCCTGTTCCATGCTCCAGCCTTGCTTAT 73
DB 1228 ATGGAATTCTACCTGCTCGTCCACAGTTGTTTATCTCATCTGATAATTCATCCATAC 1169
QY 74 TCTGAACCAAGAGATACCTACTCCCAACAAAT 106
DB 1168 CAAGAATCGCAATAATATCTTGTAACCTCTCTAT 1136

Search completed: September 15, 2005, 08:25:04
Job time : 31.7481 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 07:10:11 ; Search time 102.248 Seconds
(without alignments)
10230.248 Million cell updates/sec

Title: US-10-713-381-1_COPY_1155_1311

Perfect score: 157

Sequence: 1 cgcgtcatctcacatgscat.....cgcacatctgtcttcgaac 157

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 7389322 seqs, 333185599 residues

Total number of hits satisfying chosen parameters: 14778644

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/2/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/prodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq:*
7: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/prodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/prodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/prodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/prodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/prodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/prodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/prodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/prodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/prodata/2/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/prodata/2/pubpna/US10G_PUBCOMB.seq:*
20: /cgn2_6/prodata/2/pubpna/US10H_PUBCOMB.seq:*
21: /cgn2_6/prodata/2/pubpna/US10I_PUBCOMB.seq:*
22: /cgn2_6/prodata/2/pubpna/US10J_NEW_PUB.seq:*
23: /cgn2_6/prodata/2/pubpna/US11A_PUBCOMB.seq:*
24: /cgn2_6/prodata/2/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	157	100.0	255	20	US-10-713-381-9
2	157	100.0	1394	20	US-10-713-381-1
3	157	100.0	1394	20	US-10-713-381-2
4	146	93.0	158	20	US-10-713-381-3
5	50	31.8	50	20	US-10-713-381-5
6	40	25.5	40	20	US-10-713-381-6
7	32.2	20.5	399	19	US-10-674-124A-9081

C	8	31.6	20.1	2352	17	US-10-282-122A-16272	Sequence 36272, A
C	9	30.8	19.6	1122	19	US-10-437-963-1496	Sequence 1496, Ap
C	10	30.6	19.5	559	13	US-10-027-632-244811	Sequence 244811, Sequence 244811,
C	11	30.6	19.5	559	17	US-10-027-632-244811	Sequence 223, App
C	12	30.6	19.5	79467	18	US-10-052-482-223	Sequence 11, Appl
C	13	30.4	19.4	52001	19	US-10-317-273-11	Sequence 4, Appl
C	14	30.4	19.1	30	20	US-10-713-381-4	Sequence 170626,
C	15	30	19.1	843	13	US-10-027-632-170626	Sequence 170626,
C	16	30	19.1	843	17	US-10-027-632-170626	Sequence 170626,
C	17	29.8	19.0	236	9	US-09-783-590-10490	Sequence 57018, A
C	18	29.8	19.0	691	18	US-10-424-559-57018	Sequence 179264,
C	19	29.8	19.0	2735	13	US-10-027-632-254612	Sequence 254612,
C	20	29.8	19.0	2735	17	US-10-027-632-254612	Sequence 254612,
C	21	29.8	19.0	1223197	13	US-10-027-632-179264	Sequence 179264,
C	22	29.8	19.0	1223197	17	US-10-027-632-179264	Sequence 2154, Ap
C	23	29.6	18.9	680	19	US-10-767-701-2154	Sequence 11, Appl
C	24	29.2	18.6	1475	17	US-10-321-802-11	Sequence 1, Appl
C	25	29.2	18.6	1830121	17	US-10-329-670-1	Sequence 1, Appl
C	26	29.2	18.6	1830121	20	US-10-158-865-1	Sequence 178606, A
C	27	29.2	18.6	1830121	22	US-10-981-687-1	Sequence 222444, A
C	28	29	18.5	358	20	US-10-425-115-78606	Sequence 222444,
C	29	29	18.5	553	13	US-10-027-632-222444	Sequence 222444,
C	30	29	18.5	553	13	US-10-027-632-222444	Sequence 222444,
C	31	29	18.5	553	17	US-10-027-632-222444	Sequence 222444,
C	32	29	18.5	553	17	US-10-027-632-222444	Sequence 85365, A
C	33	29	18.5	7192	19	US-10-437-963-85365	Sequence 21721, A
C	34	28.8	18.3	732	13	US-10-027-632-21721	Sequence 21722, A
C	35	28.8	18.3	732	13	US-10-027-632-21722	Sequence 21721, A
C	36	28.8	18.3	732	17	US-10-027-632-21721	Sequence 21721, A
C	37	28.8	18.3	732	17	US-10-027-632-21722	Sequence 127179, A
C	38	28.8	18.3	733	13	US-10-027-632-127179	Sequence 127180,
C	39	28.8	18.3	733	13	US-10-027-632-127180	Sequence 127179,
C	40	28.8	18.3	733	17	US-10-027-632-127180	Sequence 53712, A
C	41	28.8	18.3	733	17	US-10-027-632-53712	Sequence 53712, A
C	42	28.8	18.3	715517	13	US-10-027-632-53712	Sequence 53712, A
C	43	28.8	18.3	715517	17	US-10-027-632-53712	Sequence 17846, A
C	44	28.6	18.2	220	20	US-10-425-115-141945	
C	45	28.6	18.2	280	20	US-10-425-115-17846	

ALIGNMENTS

RESULT 1
US-10-713-381-9
; Sequence 9, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 9
; LENGTH: 255
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-9
Query Match 100.0%; Score 157; DB 20; Length 255;
Best Local Similarity 100.0%; Pred. No. 1.7e-43;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGTCATCTCACATGSCATCTACTACTGCTTGTTCACCGCTTCGTTGTCATCTGCC 60

```

Db      15 CGGTGATCTCAACATGAGATACATGCTGTGTTCAACGGTTCGTCTTGTCCATGCTCC 74
Qy      61 AAGCCTTGCCCTATTCTGAAACCAAGAGATACCTACTCCCAACAATCCATCTTACTCAGT 120
Db      75 AAGCCTTGCCCTATTCTGAAACCAAGAGATACCTACTCCCAACAATCCATCTTACTCAGT 134
Qy      121 CAACCTTCATGCAACACGACATATGTTTCTCGAAG 157
Db      135 CAACCTTCATGCAACACGACATATGTTTCTCGAAG 171

RESULT 2
US-10-713-381-1
; Sequence 1, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-1

Query Match      100.0%; Score 157; DB 20; Length 1394;
Best Local Similarity 100.0%; Pred. No. 3.1e-43;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CGGTGATCTCAACATGAGATACATGCTGTGTTCAACGGTTCGTCTTGTCCATGCTCC 60
Db      1155 CGGTGATCTCAACATGAGATACATGCTGTGTTCAACGGTTCGTCTTGTCCATGCTCC 1214
Qy      61 AAGCCTTGCCCTATTCTGAAACCAAGAGATACCTACTCCCAACAATCCATCTTACTCAGT 120
Db      1215 AAGCCTTGCCCTATTCTGAAACCAAGAGATACCTACTCCCAACAATCCATCTTACTCAGT 1274
Qy      121 CAACCTTCATGCAACACGACATATGTTTCTCGAAG 157
Db      1275 CAACCTTCATGCAACACGACATATGTTTCTCGAAG 1311

RESULT 3
US-10-713-381-2
; Sequence 2, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0

```

```

; SEQ ID NO 2
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-2

Query Match      100.0%; Score 157; DB 20; Length 1394;
Best Local Similarity 100.0%; Pred. No. 3.1e-43;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CGGTGATCTCAACATGAGATACATGCTGTGTTCAACGGTTCGTCTTGTCCATGCTCC 60
Db      1155 CGGTGATCTCAACATGAGATACATGCTGTGTTCAACGGTTCGTCTTGTCCATGCTCC 1214
Qy      61 AAGCCTTGCCCTATTCTGAAACCAAGAGATACCTACTCCCAACAATCCATCTTACTCAGT 120
Db      1215 AAGCCTTGCCCTATTCTGAAACCAAGAGATACCTACTCCCAACAATCCATCTTACTCAGT 1274
Qy      121 CAACCTTCATGCAACACGACATATGTTTCTCGAAG 157
Db      1275 CAACCTTCATGCAACACGACATATGTTTCTCGAAG 1311

RESULT 4
US-10-713-381-3
; Sequence 3, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 158
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-3

Query Match      93.0%; Score 146; DB 20; Length 158;
Best Local Similarity 99.4%; Pred. No. 8.2e-40;
Matches 157; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy      1 CGGTGATCTCAACATGAGATACATGCTGTGTTCAACGGTTCGTCTTGTCCATGCTCC 59
Db      1 CGGTGATCTCAACATGAGATACATGCTGTGTTCAACGGTTCGTCTTGTCCATGCTCC 60
Qy      60 CAAGCTTGCCCTATTCTGAAACCAAGAGATACCTACTCCCAACAATCCATCTTACTCAGT 119
Db      61 CAAGCTTGCCCTATTCTGAAACCAAGAGATACCTACTCCCAACAATCCATCTTACTCAGT 120
Qy      120 GCACTTCATGCAACACGACATATGTTTCTCGAAG 157
Db      121 GCACTTCATGCAACACGACATATGTTTCTCGAAG 158

RESULT 5
US-10-713-381-5
; Sequence 5, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAAT, CARL W.

```

```
/ APPLICANT: HUFEMAN, GARY
/ APPLICANT: KENDALL, TIMMY L.
/ TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
/ TITLE OF INVENTION: USING SAME
/ FILE REFERENCE: 578R
/ CURRENT APPLICATION NUMBER: US/10/713,381
/ CURRENT FILING DATE: 2003-11-14
/ PRIOR APPLICATION NUMBER: 08/880,499
/ PRIOR FILING DATE: 1997-06-23
/ NUMBER OF SEQ ID NOS: 24
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 5
/ LENGTH: 50
/ TYPE: DNA
/ ORGANISM: Zea mays
US-10-713-381-5

Query Match          31.8%; Score 50; DB 20; Length 50;
Best Local Similarity 100.0%; Pred. No. 4,6e-07;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      75 CTGAACCAAGAGTACTTCTCCCAACATTCATCTTACTCATGCAAC 124
Db      1 CTGAACCAAGAGTACTTCTCCCAACATTCATCTTACTCATGCAAC 50

RESULT 6
US-10-713-381-6
/ Sequence 6, Application US/10713381
/ Publication No. US20040221331A1
/ GENERAL INFORMATION:
/ APPLICANT: ALBERTSEN, MARC C.
/ APPLICANT: FOX, TIMOTHY W.
/ APPLICANT: GARNAT, CARL W.
/ APPLICANT: HUFEMAN, GARY
/ APPLICANT: KENDALL, TIMMY L.
/ TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
/ TITLE OF INVENTION: USING SAME
/ FILE REFERENCE: 578R
/ CURRENT APPLICATION NUMBER: US/10/713,381
/ CURRENT FILING DATE: 2003-11-14
/ PRIOR APPLICATION NUMBER: 08/880,499
/ PRIOR FILING DATE: 1997-06-23
/ NUMBER OF SEQ ID NOS: 24
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 6
/ LENGTH: 40
/ TYPE: DNA
/ ORGANISM: Zea mays
US-10-713-381-6

Query Match          25.5%; Score 40; DB 20; Length 40;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      85 AGGATCTACTCTCCCAACATTCATCTTACTCATGCAAC 124
Db      1 AGGATCTACTCTCCCAACATTCATCTTACTCATGCAAC 40

RESULT 7
US-10-674-124A-9081/C
/ Sequence 9081, Application US/10674124A
/ Publication No. US20040197797A1
/ GENERAL INFORMATION:
/ APPLICANT: INOKO, HIDEKOHI
/ APPLICANT: TAMURA, GEN
/ TITLE OF INVENTION: GENE MAPPING METHOD USING MICROSATELLITE
/ TITLE OF INVENTION: GENETIC POLYMORPHISM MARKERS
/ FILE REFERENCE: ORIN-0031CIP
/ CURRENT APPLICATION NUMBER: US/10/674,124A
/ CURRENT FILING DATE: 2003-09-26
/ PRIOR APPLICATION NUMBER: 10/257,511
```

```
/ PRIOR FILING DATE: 2003-03-07
/ PRIOR APPLICATION NUMBER: PCT/JP00/07621
/ PRIOR FILING DATE: 2000-10-30
/ PRIOR APPLICATION NUMBER: JP2000-112699
/ PRIOR FILING DATE: 2000-04-13
/ PRIOR APPLICATION NUMBER: JP2002-327516
/ PRIOR FILING DATE: 2002-09-28
/ PRIOR APPLICATION NUMBER: JP2002-383869
/ PRIOR FILING DATE: 2002-12-09
/ NUMBER OF SEQ ID NOS: 27110
/ SEQ ID NO 9081
/ LENGTH: 399
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: chr5.fa.07frz.139538206
/ FEATURE:
/ OTHER INFORMATION: Located on chromosome 5
/ FEATURE:
/ OTHER INFORMATION: Distance between a terminus base of telomere on
/ OTHER INFORMATION: chromosomal short arm and 5'-terminus of this base
/ FEATURE:
/ OTHER INFORMATION: sequence : 128437212
/ OTHER INFORMATION: Distance between 3'-terminus of neighbour sequence of
/ OTHER INFORMATION: sequence listing upward to telomere on chromosomal short arm and
/ OTHER INFORMATION: 5'-terminus of this base sequence : 250186
US-10-674-124A-9081

Query Match          20.5%; Score 32.2; DB 19; Length 399;
Best Local Similarity 61.2%; Pred. No. 1.3;
Matches 52; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

OY      72 ATTCTGAACCAAGAGTACTTCTCCCAACATTCATCTTACTCATGCAACTTCGARG 131
Db      239 ATTTGAAGAAACAGTGGGCCCTCAAAAACCAACATCTTAACTTAAAAA 180

OY      132 CAACACGACATATCTTCCGAA 156
Db      179 AAAAAACACACATTTGTGATGAA 155

RESULT 8
US-10-282-122A-36272/C
/ Sequence 36272, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ CURRENT FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
```

;; PRIOR FILING DATE: 2000-10-23
;; PRIOR APPLICATION NUMBER: 60/253,625
;; PRIOR FILING DATE: 2000-11-27
;; PRIOR APPLICATION NUMBER: 60/257,931
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: 60/267,636
;; PRIOR FILING DATE: 2001-02-09
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; Remaining Prior Application data removed - See File Wrapper or PAM.
;; NUMBER OF SEQ ID NOS: 78614
;; SOFTWARE: Patent version 3.1
;; SEQ ID NO 36272
;; LENGTH: 2352
;; TYPE: DNA
;; ORGANISM: Streptococcus mutans
US-10-282-122A-36272

Query Match 20.4%; Score 31.6; DB 17; Length 2352;
Best Local Similarity 58.5%; Pred. No. 3.9;
Matches 55; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 62 AGCTTGCCTATTCTGACCAAGAGTACTCTCCCAACATCATCTTCTCATGC 121
DB 234 ATCTGACTGTGCCCAAGCCAAAGAGCAAGATCTGTAATAATCTGTCTCATCTGC 175
QY 122 AACTTCATGCAACAGCAGCATATATGTTCTCGA 155
DB 174 AATTTCAGCAAAAGACGTTCAATTTTCTCTGA 141

RESULT 9

US-10-437-963-1496/c
;; Sequence 1496, Application US/10437963
;; Publication No. US20040123343A1
;; GENERAL INFORMATION:
;; APPLICANT: La Rosa, Thomas J.
;; APPLICANT: Kovalic, David K.
;; APPLICANT: Zhou, Yihua
;; APPLICANT: Cao, Yongwei
;; APPLICANT: Wu, Wei
;; APPLICANT: Boukharov, Andrey A.
;; APPLICANT: Barbazuk, Brad
;; APPLICANT: Li, Ping
;; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
;; FILE REFERENCE: 38-21(53221)B
;; CURRENT APPLICATION NUMBER: US/10/437,963
;; CURRENT FILING DATE: 2003-05-14
;; NUMBER OF SEQ ID NOS: 204966
;; SEQ ID NO 1496
;; LENGTH: 1122
;; TYPE: DNA
;; ORGANISM: Oryza sativa
;; FEATURE:
;; OTHER INFORMATION: Clone ID: PAT_MRT4530_101357C.1
US-10-437-963-1496

Query Match 19.6%; Score 30.8; DB 19; Length 1122;
Best Local Similarity 57.1%; Pred. No. 5.6;
Matches 56; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 60 CAAGCTTGCTATTCTGAACCAAGAGTACTCTCCCAACATCATCTTACTCAT 119
DB 313 CAGCTCGCATCATTTGCTATCCGGGAGTCACTAAATCAACAAAGCCATATTCATTCT 254
QY 120 GCAACTTCATGCAACAGCAGCATATGTTTCTGAAC 157
DB 253 TCAGATCCACAATTCGACCCCTCCCATATCTGCTGAAC 216

RESULT 10
US-10-027-632-244811

;; Sequence 244811, Application US/10027632
;; Publication No. US20020198371A1
;; GENERAL INFORMATION:
;; APPLICANT: Wang, David G.
;; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;; TITLE OF INVENTION: Polymorphisms in the Human Genome
;; FILE REFERENCE: 108827.129
;; CURRENT APPLICATION NUMBER: US/10/027,632
;; CURRENT FILING DATE: 2002-04-30
;; PRIOR APPLICATION NUMBER: US 60/218,006
;; PRIOR FILING DATE: 2000-07-12
;; PRIOR APPLICATION NUMBER: US 60/198,676
;; PRIOR FILING DATE: 2000-04-20
;; PRIOR APPLICATION NUMBER: US 60/193,483
;; PRIOR FILING DATE: 2000-03-29
;; PRIOR APPLICATION NUMBER: US 60/185,218
;; PRIOR FILING DATE: 2000-02-24
;; PRIOR APPLICATION NUMBER: US 60/167,363
;; PRIOR FILING DATE: 1999-11-23
;; PRIOR APPLICATION NUMBER: US 60/156,358
;; PRIOR FILING DATE: 1999-09-28
;; PRIOR APPLICATION NUMBER: US 60/146,002
;; PRIOR FILING DATE: 1999-08-09
;; NUMBER OF SEQ ID NOS: 325720
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 244811
;; LENGTH: 559
;; TYPE: DNA
;; ORGANISM: Human
US-10-027-632-244811

Query Match 19.5%; Score 30.6; DB 13; Length 559;
Best Local Similarity 51.9%; Pred. No. 5.1;
Matches 69; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 24 ACATGCTGTTCACCGTTCCTGTTTCATGTGTCACAGCCTGCTATTCTGACCA 83
DB 380 ACATGCAATTTTCACTTTCGATTTCGATTCCTGCTGAGGTACTTTTCTAGATTCA 439
QY 84 GAGGATTAATCTCTCCCAACCAATCATCTTACTCATGAACTTCATGCAACAGCACA 143
DB 440 TCTGCTCCCTACTTTTAAAGATTCATGAGAGTCTCTACAAATCCATAGCACACTGAAC 499
QY 144 TATGTTCTCTGAA 156
DB 500 ATTTCTTCCCA 512

RESULT 11

US-10-027-632-244811
;; Sequence 244811, Application US/10027632
;; Publication No. US20030204075A9
;; GENERAL INFORMATION:
;; APPLICANT: Wang, David G.
;; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;; TITLE OF INVENTION: Polymorphisms in the Human Genome
;; FILE REFERENCE: 108827.129
;; CURRENT APPLICATION NUMBER: US/10/027,632
;; CURRENT FILING DATE: 2002-04-30
;; PRIOR APPLICATION NUMBER: US 60/218,006
;; PRIOR FILING DATE: 2000-07-12
;; PRIOR APPLICATION NUMBER: US 60/198,676
;; PRIOR FILING DATE: 2000-04-20
;; PRIOR APPLICATION NUMBER: US 60/193,483
;; PRIOR FILING DATE: 2000-03-29
;; PRIOR APPLICATION NUMBER: US 60/185,218
;; PRIOR FILING DATE: 2000-02-24
;; PRIOR APPLICATION NUMBER: US 60/167,363
;; PRIOR FILING DATE: 1999-11-23
;; PRIOR APPLICATION NUMBER: US 60/156,358
;; PRIOR FILING DATE: 1999-09-28
;; PRIOR APPLICATION NUMBER: US 60/146,002
;; PRIOR FILING DATE: 1999-08-09

```
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 244811
; LENGTH: 359
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-244811

Query Match      19.5%; Score 30.6; DB 17; Length 559;
Best Local Similarity 51.9%; Pred. No. 5.1;
Matches 69; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 24 ACATGCTGTTCACCGTGTGCTTGTTCATGCGCAAGCCTGCGATTGCAACCA 83
    |||||
DB 380 ACATGCTGTTCATCTTTTCATTTGCTCATCTTGTCTGAGGACTTTTCTAGATTCA 439

QY 84 GAGGATACCTACTCCCAACATCCATCTTACTGATGCACTTCATGCAACAGCACA 143
    |||||
DB 440 TCTGCTCCCTTAAAGATTTCATGGAAGTCTTCAAAATCCCATGACACTGAC 499

QY 144 TATGTTCTCGAA 156
    |||||
DB 500 ATTTCTTCCCA 512

RESULT 12
US-10-052-482-223
; Sequence 223, Application US/10052482
; Publication No. US20040072264A1
; GENERAL INFORMATION:
; APPLICANT: Morriss, David
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: A-71087/RMS/DCF
; CURRENT APPLICATION NUMBER: US/10/052,482
; PRIOR FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 223
; LENGTH: 79467
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5502)..(5521)
; OTHER INFORMATION: "n" at positions 5502 to 5521 can be any base
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4095)..(4369)
; OTHER INFORMATION: "n" at positions 4095 to 4369 can be any base
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10089)..(10620)
; OTHER INFORMATION: "n" at positions 10089 to 10620 can be any base
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (13273)..(13370)
; OTHER INFORMATION: "n" at positions 13273 to 13370 can be any base
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (20762)..(20781)
; OTHER INFORMATION: "n" at positions 20762 to 20781 can be any base
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (30916)..(30916)
; OTHER INFORMATION: "n" at positions 30916 can be any base
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (46579)..(46772)
```

```
; OTHER INFORMATION: "n" at positions 46759 to 46772 can be any base
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (77382)..(77401)
; OTHER INFORMATION: "n" at positions 77372 to 77401 can be any base
US-10-052-482-223

Query Match      19.5%; Score 30.6; DB 18; Length 79467;
Best Local Similarity 56.4%; Pred. No. 31;
Matches 57; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 51 TCCATGCTCCAGCCTTGCTTATTCGACCAAGAGATACCTACTCCCAACATTCAT 110
    |||||
DB 53779 TCCATGCTCCAGCCTTGCTTATTCGACCAAGAGATACCTACTCCCAACATTCAT 110

QY 111 CTACTGATGCACTTCATGCAACAGCAGCATATGTTTC 151
    |||||
DB 53839 CACACACACACACACACACACACACAGCTTGCGTC 53879

RESULT 13
US-10-317-273-11
; Sequence 11, Application US/10317273
; Publication No. US20040110158A1
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Tamara Balac Stipes
; TITLE OF INVENTION: MODULATION OF MITOCHONDRIAL RIBOSOMAL PROTEIN L13 EXPRESSION
; FILE REFERENCE: RTS-0478
; CURRENT APPLICATION NUMBER: US/10/317,273
; CURRENT FILING DATE: 2002-12-10
; NUMBER OF SEQ ID NOS: 159
; SEQ ID NO 11
; LENGTH: 52001
; TYPE: DNA
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 11121, 11122, 11123, 11124, 11125, 11126, 11127, 11128, 11129,
; LOCATION: 11130, 11131, 11132, 11133, 11134, 11135, 11136, 11137,
; LOCATION: 11138, 11139, 11140, 11141, 11142, 11143, 11144, 11145,
; LOCATION: 11146, 11147, 11148, 11149, 11150, 11151, 11152, 11153
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 11154, 11155, 11156, 11157, 11158, 11159, 11160, 11161, 11162,
; LOCATION: 11163, 11164, 11165, 11166, 11167, 11168, 11169, 11170,
; LOCATION: 11171, 11172, 11173, 11174, 11175, 11176, 11177, 11178,
; LOCATION: 11179, 11180, 11181, 11182, 11183, 11184, 11185, 11186
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 11187, 11188, 11189, 11190, 11191, 11192, 11193, 11194, 11195,
; LOCATION: 11196, 11197, 11198, 11199, 11200, 11201, 11202, 11203,
; LOCATION: 11204, 11205, 11206, 11207, 11208, 11209, 11210, 11211,
; LOCATION: 11212, 11213, 11214, 11215, 11216, 11217, 11218, 11219
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 11220
; OTHER INFORMATION: n = A,T,C or G
US-10-317-273-11

Query Match      19.4%; Score 30.4; DB 19; Length 52001;
Best Local Similarity 61.2%; Pred. No. 31;
Matches 49; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 50 TTCATGCTCCAGCCTTGCTTATTCGACCAAGAGATACCTACTCCCAACATTCAT 109
    |||||
DB 20055 TTCCTTTTCTAAAGTAATTCATGAAAAAGATATCAATTTTAAACTACCA 20114

QY 110 TCTTACTCATGCACTTCCA 129
    |||||
```

Db 20115 TCTTCTCAGCAAACTTCTA 20134

RESULT 14

US-10-713-381-4
Sequence 4, Application US/10713381
Publication No. US20040221331A1
GENERAL INFORMATION:
APPLICANT: ALBERTSEN, MARC C.
APPLICANT: FOX, TIMOTHY W.
APPLICANT: GARNAT, CARL W.
APPLICANT: HUFEMAN, GARY
APPLICANT: KENDALL, TIMMY L.
TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
FILE REFERENCE: 578R
CURRENT APPLICATION NUMBER: US/10/713,381
CURRENT FILING DATE: 2003-11-14
PRIOR APPLICATION NUMBER: 08/880,499
PRIOR FILING DATE: 1997-06-23
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 30
TYPE: DNA
ORGANISM: Zea mays
US-10-713-381-4

Query Match 19.1%; Score 30; DB 20; Length 30;
Best Local Similarity 100.0%; Pred. No. 2.8;

Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 CATGCTGTTCACCGTTCGTTGTTCCA 54
Db 1 CATGCTGTTCACCGTTCGTTGTTCCA 30

RESULT 15

US-10-027-632-170626
Sequence 170626, Application US/10027632
Publication No. US20020198371A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/195,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 170626
LENGTH: 843
TYPE: DNA
ORGANISM: Human
US-10-027-632-170626

Query Match 19.1%; Score 30; DB 13; Length 843;

Best Local Similarity 51.5%; Pred. No. 9.5;

Matches 69; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 6 CATTCACATGAGCATACTACATGTTGTTCAACCGTTCGTTGTTCCATGTCACAGCC 65
Db 552 CAGCACCCAGGCCAGCTCAAGACAACGTCACTCTCTGCGTGGGCCCCACATCCATGCC 611
QY 66 TTGCCTATTCTGAAACCAAGAGATACCTATCCCAACATGCATCTTACTGATGCACT 125
Db 612 TTGCATGTGCTGGAAGATGGGGGCTTCGCTCCCTGCTCATCAACGCTGGGAGGCGCTCC 671
QY 126 TCCATGCAACACG 139
Db 672 TGCATCCGAACACG 685

Search completed: September 15, 2005, 20:45:38
Job time : 108.248 secs


```

QY      61  AACCTTGCTATTTGAAACCAAGAGATCTACTCTCCAAATCATCTTACTCAG 120
DB      361  AAGCTTGCTATTTGAAACCAAGAGATCTACTCTCCAAATCATCTTACTCAG 302
QY      121  CAACCTTCATGCAACACGACATATGTTCTCTGAAC 157
DB      301  CAACCTTCATGCAACACGACATATGTTCTCTGAAC 265

RESULT 2
LOCUS   CG224225
DEFINITION  915 bp DNA linear GSS 22-AUG-2003
GSS: OGIAG08TV ZM 0.7-1.5 KB zea maye genomic clone ZMMBMA0716B15,
genomic survey sequence.
ACCESSION  CG224225
VERSION    CG224225
KEYWORDS   GI:34124113
SOURCE     GSS.
ORGANISM   Zea mays
            Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
            1 (bases 1 to 915)
REFERENCE  WhiteIaw,C.A., Quackenbush,D., Van Aken,S., Uteerback,T.,
AUTHORS    Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
TITLE      Consortium for Maize Genomics
JOURNAL    Unpublished (2002)
COMMENT    Other GSSs: OGIAG08TH
            Contact: Cathy WhiteIaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whiteIaw@tigr.org
            Seq primer: TF
            Classes: sheared ends.

FEATURES
source     Location/Qualifiers
            1..915
            /organism="Zea mays"
            /mol_type="genomic DNA"
            /strain="B73"
            /db_xref="taxon:4577"
            /clone="ZMMBMA0716B15"
            /note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
            methylation filtered genomic DNA library"

ORIGIN
Query Match 100.0%; Score 157; DB 9; Length 915;
Best Local Similarity 100.0%; Pred. No. 1.9e-39;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  CGTGTATCTACATGGAATCTACTACTGCTTGTTCACCGTTCGTTCCATCGTCC 60
DB      534  CGTGTATCTACATGGAATCTACTACTGCTTGTTCACCGTTCGTTCCATCGTCC 593
QY      61  AAGCTTGCTATTTGAAACCAAGAGATCTACTCTCCAAATCATCTTACTCAG 120
DB      594  AAGCTTGCTATTTGAAACCAAGAGATCTACTCTCCAAATCATCTTACTCAG 653
QY      121  CAACCTTCATGCAACACGACATATGTTCTCTGAAC 157
DB      654  CAACCTTCATGCAACACGACATATGTTCTCTGAAC 690

RESULT 3
LOCUS   CC656933
DEFINITION  963 bp DNA linear GSS 19-JUN-2003
GSS: OGMDD20TM ZM 0.7-1.5 KB zea maye genomic clone ZMMBMA0554D15,
genomic survey sequence.
ACCESSION  CC656933

```

```

VERSION   CC656933.1  GI:32060225
KEYWORDS  GSS.
SOURCE    Zea mays
ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
            1 (bases 1 to 963)
REFERENCE  WhiteIaw,C.A., Quackenbush,D., Van Aken,S., Uteerback,T.,
AUTHORS    Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
TITLE      Consortium for Maize Genomics
JOURNAL    Unpublished (2002)
COMMENT    Other GSSs: OGMDD20TM
            Contact: Cathy WhiteIaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whiteIaw@tigr.org
            Seq primer: TR
            Classes: sheared ends.

FEATURES
source     Location/Qualifiers
            1..963
            /organism="Zea mays"
            /mol_type="genomic DNA"
            /strain="B73"
            /db_xref="taxon:4577"
            /clone="ZMMBMA0554D15"
            /note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
            methylation filtered genomic DNA library"

ORIGIN
Query Match 100.0%; Score 157; DB 9; Length 963;
Best Local Similarity 100.0%; Pred. No. 2e-39;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  CGTGTATCTACATGGAATCTACTACTGCTTGTTCACCGTTCGTTCCATCGTCC 60
DB      752  CGTGTATCTACATGGAATCTACTACTGCTTGTTCACCGTTCGTTCCATCGTCC 811
QY      61  AAGCTTGCTATTTGAAACCAAGAGATCTACTCTCCAAATCATCTTACTCAG 120
DB      812  AAGCTTGCTATTTGAAACCAAGAGATCTACTCTCCAAATCATCTTACTCAG 871
QY      121  CAACCTTCATGCAACACGACATATGTTCTCTGAAC 157
DB      872  CAACCTTCATGCAACACGACATATGTTCTCTGAAC 908

RESULT 4
LOCUS   CE588104/c
DEFINITION  636 bp DNA linear GSS 28-SEP-2003
GSS: tigr-gss-dog-17000366359543 dog library Canis familiaris genomic,
genomic survey sequence.
ACCESSION  CE588104
VERSION    CE588104.1  GI:36904885
KEYWORDS   GSS.
SOURCE     Canis familiaris (dog)
ORGANISM   Canis familiaris
            Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
            Eukaryota; Metazoa; Chordata; Carnivora; Fissipedia; Canidae; Canis.
            1 (bases 1 to 636)
REFERENCE  Kirkness,B.F., Batina,V., Halpern,A.L., Levy,S., Remington,K.,
AUTHORS    Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
            Venter,J.C.
TITLE      The dog genome: survey sequencing and comparative analysis
JOURNAL    Science 301 (5641), 1898-1903 (2003)
MEDLINE    22875432
PUBMED     14512627
COMMENT    Contact: Kirkness EF
            The Institute for Genomic Research

```


LOCUS BB505306 700 bp mRNA linear EST 25-OCT-2001
DEFINITION BB505306 RIKEN full-length enriched, 10 days lactation, adult female mammary gland Mus musculus cDNA clone D730002G06, mRNA sequence.
ACCESSION BB505306
VERSION BB505306.2 GI:16442791
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 700)
AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,M., Koyu,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
Unpublished (2001)
On Jul 27, 2000 this sequence version replaced gi:9514268.
Contact: Yoshinide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.jp, URL:http://genome.gsc.riken.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Komu,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Wataniki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kikura,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Komu,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and Hayashizaki,Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

FEATURES

source

1..700
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="D730002G06"
/sex="female"
/tissue_type="mammary gland"
/dev_stage="10 days lactation, adult"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 10 days lactation, adult female mammary gland"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia"

ORIGIN

Query Match 21.7%; Score 34; DB 2; Length 700;
Best Local Similarity 54.9%; Pred. No. 6.8;
Matches 67; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
QY 27 TCGTGTCAACCGTTCGTCCTGTCATCGCCAAAGCCTTGCTGAACCAAGG 86
DB 195 TTTCTTTTCTCCCTCCCTCCCTGATTTCTTTGACTCTACCTTTTGAGCTTTBAAGT 254
QY 87 GATTACTACTCCCAACAATCCATCTTAATGCAACTTCATGCAACGACGACATAT 146
DB 255 GTTCTAGATTCCCAAGAAAGCATTGTGACATGACATGTAACATATCAACCTGAAAT 314
QY 147 GT 148
DB 315 GT 316

RESULT 8

LOCUS CL026582 1677 bp DNA linear GSS 31-DEC-2003
DEFINITION CH216-23G24 Sp6.1 CH216 Xenopus tropicalis genomic clone
ACCESSION CL026582
VERSION CL026582.1 GI:40470443
KEYWORDS GSS.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae; Xenopodinae; Xenopus; Silurana.

REFERENCE 1 (bases 1 to 1677)
AUTHORS Krenitschki,C., Carter,J., McPherson,J., Warren,W., Graves,T., Mardis,E. and Wilson,R.
A physical map of the xenopus tropicalis genome
Unpublished (2003)
COMMENT Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 175000 Std Error: 0.00
Seq primer: Sp6 ATTAGGTGACACTATAG
Class: BAC ends
High quality sequence start: 125
High quality sequence stop: 217.
Location/Qualifiers

FEATURES

source

1..1677
Location/Qualifiers
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/strain="Nigerian frog"
/db_xref="taxon:8364"
/clone="CH216-23G24"
/sex="male"
/cell_line="Stock 248 F7A2, inbred N7"
/clone_lib="CH216"
/note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis BAC library"

ORIGIN

Query Match	21.4%	Score 33.6	DB 9	Length 1677
Best Local Similarity	53.9%	Pred. No. 12		
Matches 69	Conservative 0	Mismatches 59	Indels 0	Gaps 0
Ov	3	TGTGATCTCACATGCGCATACATGCTGTGTCACCGCTGCTGTGTCATGCTGCA	62	
Db	769	TGTGATCTCACATGCGCATACATGCTGTGTCACCGCTGCTGTGTCATGCTGCA	828	
Ov	63	GCCTTGCTTATTCGACCAAGAGATACCTTACCCCAACATCATCTTACTGATCA	122	
Db	829	TCATCAACATCTTCACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	888	
Ov	123	ACTTCAT 130		
Db	889	TCTTCAT 896		
RESULT 9				
LOCUS	A1744861	440 bp	mRNA	linear
DEFINITION	U16B06.X1 NCI_CGAP_Ov23 Homo sapiens cDNA clone IMAGE:2218451 3'			
ACCESSION	A1744861			
VERSION	A1744861.1	GI:5113149		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
AUTHORS	1 (bases 1 to 440)			
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .			
JOURNAL	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Unpublished (1997)			
COMMENT	Contact: Robert Strausberg, Ph.D. Email: rgs@bbs-remail.nih.gov Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R. Emert-Buck, M.D., Ph.D. cDNA Library Preparation: Life Technologies, Inc. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bdtp/image/image.html Seq primer: -40UP from GIBCO High quality sequence stop: 411. Location/Qualifiers			
FEATURES				
Source	1..440			
	/organism="Homo sapiens"			
	/mol_type="mRNA"			
	/db_xref="taxon:9606"			
	/clone="IMAGE:2218451"			
	/tissue_type="tumor, 5 pooled (see description)"			
	/lab_host="DH10B"			
	/clone_lib="NCI_CGAP_Ov23"			
	/note="Organ: ovary; Vector: pCMV-SPORT6; site_1: SalI; site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.35 kb. Tumor types include: mixed Mullerian tumor, papillary serous, clear cell, spindle cell. All are primary tumors, metastasis positive. Life Technologies catalog #: 11534-013"			
ORIGIN				
Query Match	20.9%	Score 32.8	DB 1	Length 440
Best Local Similarity	56.5%	Pred. No. 14		
Matches 61	Conservative 0	Mismatches 47	Indels 0	Gaps 0
Ov	46	CTGTTCATCGTCACACCTTGCTTATTCGACCAAGAGATACCTTACCCAAACA	105	
Db	259	CTTGGCCAGCAACTATATATATATATCTTCCTCCCAAGCAACCTCTTATTT	318	
Ov	106	TTCATCTTACTGACCACTTCATGCAACGACGACATATGTTCC	153	

DB	319	TCGATCTTAATTAAGACGGTCGATGTCTGCTGATATCTGTTCTCT	366
RESULT 10			
CR575123			
LOCUS			
DEFINITION	CR575123 XGC-tailbud-head Xenopus tropicalis cDNA clone THdA024m23		
ACCESSION	CR575123		
VERSION	CR575123.1		
KEYWORDS	5', mRNA sequence.		
SOURCE			
ORGANISM	Xenopus tropicalis (western clawed frog)		
REFERENCE			
AUTHORS	Croning, M.D.R., Ashurst, J.L., Taylor, R., Garrett, N. and Rogers, J.		
TITLE	Sanger Xenopus tropicalis EST project 2001 (2004)		
JOURNAL	Unpublished (2004)		
COMMENT	Contact: Croning MDR Sanger Institute Hinxton, Cambridgeshire, CB10 1SA, UK Email: trop@sanger.ac.uk Sanger Xenopus tropicalis EST project 2001 TROPICALIS_SEQUENCE_ID: THdA024m23.p1bSP6 This sequence is from a Xenopus Gene Collection (XGC) library constructed by Nigel Garrett. Seq primer: SP6.		
FEATURES			
source	Location/Qualifiers 1..804 /organism="Xenopus tropicalis" /mol_type="mRNA" /db_xref="taxon:8364" /clone="THdA024m23" /dev_stage="tailbud head (stage 28-30)" /lab_host="Baccharichia coli DH10B." /note="Vector: pCS107; Site 1: EcoRI, Site 2: NotI; cDNA was oligo dt primed from 5ug of poly A+ RNA from tailbud head. EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the 5' end and NotI at the 3' end."		
ORIGIN			
Query Match	20.8%; Score 32.6; DB 7; Length 804;		
Best Local Similarity	63.3%; Pred. No. 20;		
Matches	50; Conservative 0; Mismatches 29; Indels 0; Gaps 0;		
OY	74 TCTGAACCAAGAGANTACTACTCCCAACATCATCTTACTGATGCAACTTCATGCA 133		
DB	480 TCTGAATATATGAAGAAGCATCTCCATCTAGACTCATTTATGCAAGTAATTTACATTTT 539		
OY	134 AACACGACATATGTTCC 152		
DB	540 AAAAAGTATTTTTC 558		
RESULT 11			
CNS003FV/c			
LOCUS			
DEFINITION	CNS003FV 1101 bp DNA linear GSS 03-JUN-1999		
ACCESSION	Drosophila melanogaster genome survey sequence TET3 end of BAC #		
VERSION	BACR08003 of Rpci-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.		
KEYWORDS	AT064657		
SOURCE	AT064657.1 GI:4942009		
ORGANISM	GSS.		
REFERENCE	Drosophila melanogaster (fruit fly)		
AUTHORS	Drosophila melanogaster		
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;		
	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
	Ephydroidea; Drosophilidae; Drosophila.		
	1 (base 1 to 1101)		
	Genoscope.		

```

FEATURES
  source
    location/Qualifiers
      1..1101
        /organism="Drosophila melanogaster"
        /mol_type="genomic DNA"
        /db_xref="taxon:7227"
        /clone="BACRO8003"
        /clone_1b="RPCI-98"
        /note="end : TET3"

```

QY 29 CTGTTCACCGTTCGTTGTCATGCGTCCAGCGCTTGCTATTTCGAACCAAGAGA 88
Db 997 CBNNTKDAAYDKDQDATKBSKCYCMAAIVSCCARDNWTMAACHDKKYCMCGMWHNAHA 938
QY 89 TACTTACTCCCAACCAATCCATCTTACTATGCAACTTCATGCAACACCCAC 142
Db 937 WDNKKKXCCMAAAMARBCKRDKKTTTWMKTXCYCYCAKAGRYKSCCC 884

ACCESSION	BE690384
VERSION	BE690384.1
KEYWORDS	GI:10078008
SOURCE	EST.
ORGANISM	Mus musculus (house mouse)
	Mus musculus

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

COMMENT **Contact: Robert Strausberg, Ph.D.**

Seq primer: -40RP from Gibco
High quality sequence stop:

FEATURES	Location/Qualifiers
SOURCE	1. ,347

```

/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:3466991"
/sex="female (lactating)"

```

	ORIGIN
Query Match	20.6%; Score 32.4; DB 2; Length 347;
Best Local Similarity	54.1%; Pred. No. 18;
Matches 66; Conservative	0; Mismatches 56; Indels 0; Gaps 0;

87 GATACCTACTCCCAAGCAATTCATTCTTACTCATGCAACCTTCATGCAACACGACATAT 146
 208 GTTCTTAAGTTCCTCCAAAGACGATTTTGGCATCTACAGTCAACATCATCTTAACGCCCTGTAAT 267

QY	147	GT	148
DB	268	GT	269

ACCESSION	AJ746802	
VERSION	AJ746802.1	GI:49916882

KEYWORDS	EST.
SOURCE	Sub scrofa (pig)

ORGANISM	Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.	
REFERENCE	
1 (bases 1 to 483)	

TITLE	Development of a porcine cDNA microarray
JOURNAL	Unpublished (2004)
COMMENT	Contact: Hopwood PA

Dept. of Preclinical Veterinary Sciences
Royal School for Veterinary Studies
Summerhall, Edinburgh, EH9 1QH, UNITED KINGDOM
Sequencing was performed by ARK genomics. This clone is available
from ARK-Genomics, Roslin Institute, Roslin, Midlothian EH25 9PS,
UK. See www.ark-genomics.org or contact info@arkgenomics.org.

FEATURES	Location/Quali
source	1. .483

```

/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="Caxon:3823"
/clone="F C0001825C D07"
/tissue_type="lung"
/cell_type="macrohage"
/clone_1ib="forward - stimulated minus unstimulated
macrohage"

```

ORIGIN

Query Match	20.6%	Score 32.4	DB 1	Length 483
Best Local Similarity	55.3%	Pred. No. 20		
Matches 63	0	Mismatches 51	Indels 0	Gaps 0

QY	21	ACTACATGCTTGTTCACCGCTTCGTTCTGTTCCATCGTCCAAAGCCTTCCTATTTCGAC	80
Db	371	ACACATCTATGTTAAAGCTTTTTTTTTTTTAAAAAAGCCCAATATATTTCTTAC	312

QY 81 CAGAGGATACCTACCTCCCAACATCCATCTTACTGATGCAACTTCGATGCAA 134
DB 311 ACAGTAGAGTCCACCTACCTATCAATTCAGTTAAATGAATATTCACAGGAA 258

RESULT 14
LOCUS AW990724/c 497 bp mRNA linear EST 02-JUN-2000
DEFINITION u010b11.x1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone
ACCESSION AW990724
VERSION AW990724
KEYWORDS AW990724.1 GI:8186451
SOURCE EST.
ORGANISM Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 497)
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Other ESTs: u010b11.y1
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:937793
High quality sequence stop: 459.
Location/Qualifiers
1..497
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:1510941"
/sex="female (lactating)"
/tissue_type="mammary gland"
/lab_host="DH10B"
/lab_lib="Soares_mammary_gland_NMLMG"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from mammary
gland tissue from a lactating female, and was then primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pT73 vector. Library is normalized. Library
was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN
Query Match 20.6%; Score 32.4; DB 2; Length 497;
Best Local Similarity 54.1%; Pred. No. 20;
Matches 66; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 27 TGCTGTCAACGCTGCTTGTTCATGCTCCAGCCTTGCTTCTGAACCAAG 86
DB 490 TTTCTTTTCACCTCCTCTGATTTCTTGATCTTCACCTTATTTGGCCTTTAAGT 431

QY 87 GATACCTACTCCCAACATCCATCTTACTGATGCAACTTCGATGCAACAGCATAT 146
DB 430 GTTCTAGATTCCCAAGAGCCATTTTGACCTAGCTCAACATCTACTACACCTGAAT 371

QY 147 GT 148
DB 370 GT 369

RESULT 15
BE632212/c 498 bp mRNA linear EST 25-AUG-2000
LOCUS u011f01.x1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone
DEFINITION IMAGE:3371641 3', mRNA sequence.
ACCESSION BE632212

VERSION BE632212.1 GI:9914830
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 498)
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Other ESTs: u011f01.y2
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:1081245
High quality sequence stop: 459.
Location/Qualifiers
1..498
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:3371641"
/sex="female (lactating)"
/tissue_type="mammary gland"
/lab_host="DH10B"
/lab_lib="Soares_mammary_gland_NMLMG"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from mammary
gland tissue from a lactating female, and was then primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pT73 vector. Library is normalized. Library
was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN
Query Match 20.6%; Score 32.4; DB 2; Length 498;
Best Local Similarity 54.1%; Pred. No. 20;
Matches 66; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 27 TGCTGTCAACGCTGCTTGTTCATGCTCCAGCCTTGCTTCTGAACCAAG 86
DB 491 TTTCTTTTCACCTCCTCTGATTTCTTGATCTTCACCTTATTTGGCCTTTAAGT 432

QY 87 GATACCTACTCCCAACATCCATCTTACTGATGCAACTTCGATGCAACAGCATAT 146
DB 431 GTTCTAGATTCCCAAGAGCCATTTTGACCTAGCTCAACATCTACTACACCTGAAT 372

QY 147 GT 148
DB 371 GT 370

Search completed: September 16, 2005, 08:08:25
Job time: 496.661 secs

This Page Blank (uspto)

REFERENCE
AUTHORS
TITLE
JOURNAL

1
Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
Patent: WO 0160997-A 9 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)

FEATURES
source
1..1394
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN

Query Match 100.0%; Score 30; DB 6; Length 255;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CATGCTGTTCACCGTTCGCTTGTTCCA 30
1179 CATGCTGTTCACCGTTCGCTTGTTCCA 68

RESULT 3
AX224394 1394 bp DNA linear PAT 10-SEP-2001
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AX224394
Sequence 1 from Patent WO0160997.
AX224394
AX224394.1 GI:15554636
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
Clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
AUTHORS
TITLE
JOURNAL

1
Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
Patent: WO 0160997-A 1 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)

FEATURES
source
1..1394
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN

Query Match 100.0%; Score 30; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CATGCTGTTCACCGTTCGCTTGTTCCA 30
1179 CATGCTGTTCACCGTTCGCTTGTTCCA 1208

RESULT 4
AX224395 1394 bp DNA linear PAT 10-SEP-2001
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AX224395
Sequence 2 from Patent WO0160997.
AX224395
AX224395.1 GI:15554637
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
Clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
AUTHORS
TITLE
JOURNAL

1
Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
Patent: WO 0160997-A 2 23-AUG-2001;

FEATURES
source
1..1394
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN

Query Match 100.0%; Score 30; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CATGCTGTTCACCGTTCGCTTGTTCCA 30
1179 CATGCTGTTCACCGTTCGCTTGTTCCA 1208

RESULT 5
BD062176 1394 bp DNA linear PAT 27-AUG-2002
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BD062176
Male tissue-preferred regulatory region and method of using same.
BD062176
BD062176.1 GI:22607781
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL

1
Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G.A. and Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
Patent: JP 2001520523-A 1 30-OCT-2001;
PIONEER HI BRED INTERNATIONAL INC
PN JP 2001520523-A/1
PD 30-OCT-2001
PF 19-JUN-1998 JP 1999504910
PR 23-JUN-1997 US 08/880499
PI MARC C ALBERTSEN,TIMOTHY W FOX,CARL W GARNAAT,GARY A HUFFMAN,
PI TIMMY L KENDALL
PC C12N15/62,C12N15/29,C12N9/24,C12N9/22,C12N9/10,C12N9/00 PC
,C07K14/34,C12Q1/68,
PC A01H5/00
CC Strandedness: Single;
CC Topology: Linear;
FH key Location/Qualifiers.

FEATURES
source
1..1394
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 30; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CATGCTGTTCACCGTTCGCTTGTTCCA 30
1179 CATGCTGTTCACCGTTCGCTTGTTCCA 1208

RESULT 6
BD062177 1394 bp DNA linear PAT 27-AUG-2002
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BD062177
Male tissue-preferred regulatory region and method of using same.
BD062177
BD062177.1 GI:22607782
JP 2001520523-A/2.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 1394)
AUTHORS Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G.A. and Kendall,T.L.
TITLE Male tissue-preferred regulatory region and method of using same
JOURNAL Patent: JP 2001520523-A 2 30-OCT-2001;
PIONEER HI BRED INTERNATIONAL INC
COMMENT PN JP 2001520523-A/2
PD 30-OCT-2001
PR 19-JUN-1998 JP 1999504910
PR 23-JUN-1997 US 08/880499
PI MARC C ALBERTSEN, TIMOTHY W FOX, CARL W GARNAA, GARY A HUFFMAN,
PI TIMMY L KENDALL
PC C12N15/82, C12N15/29, C12N9/24, C12N9/22, C12N9/10, C12N9/00 PC
PC C07K14/34, C12Q1/68,
PC A01H5/00
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers.
FEATURES
source 1..1394
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 30; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CATGCTTGTTCAACCGTTCGCTGTTGCCA 30
1179 CATGCTTGTTCAACCGTTCGCTGTTGCCA 1208
Db
RESULT 7 3343 bp DNA linear PLN 12-MAY-2001
LOCUS AF360356
DEFINITION Zea mays male fertility protein (Me45) gene, complete cds.
ACCESSION AF360356
VERSION AF360356.1 GI:14028756
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 3343)
AUTHORS Fox,T.W., Trimnell,M.R. and Albertsen,M.C.
TITLE Cloning of Me45, a gene required for male fertility from Zea mays
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3343)
AUTHORS Fox,T.W., Trimnell,M.R. and Albertsen,M.C.
TITLE Direct Submision
JOURNAL Submitted (13-MAR-2001) Trait and Technology Development, Pioneer
HI-Bred Intl. Inc., 7300 N.W. 62nd Ave., P.O. Box 1004, Johnston,
IA 50131-1004, USA
FEATURES
source Location/Qualifiers
1..3343
/organism="Zea mays"
/mol_type="genomic DNA"
/culivate="B73"
/db_xref="taxon:4577"
/chromosome="9L"
/gene="Me45"
join(<1392..1768,1898..2182,2280..2447,2534..>2942)
/product="male fertility protein"
join(1392..1768,1898..2182,2280..2447,2534..2942)
/gene="Me45"
/codon_start=1
/product="male fertility protein"
/protein_id="AK52489.1"

/db_xref="GI:14028757"
/translation="MEKRIQMRREGDGI VOYPHLFPALALALLVADPEGLSPLAEV
DTPVAKHEILAPYGEWGSWPRDNASKLRGRLEFGEVGPESITIEDLOGKPYAGLA
DQRVVRMGEAEAGWETPAVMNPDWSEEVANVSTTRQHKKEEFCGRPLGRFGE
TSELVYADAYVGLVMVGOOGVASSVAREADDPTRFANDLDVHRGVSFFDTSMRY
SKRDHLNILLBEGEGRLRLRYDPETSGVAVLVKGLVFPYGVQVISEDHQLFSETNC
RLMRVWLBEGPRAGEVFPANLRPGPDNVASNGRGQWVAIDCCRPDAQDVPAKRPMLR
TLYKRPPLSLKVLTKAARMRTVVALLDDEGRVVEVLEDRGHEVWKLVSEVREVGKR
LWIGTVAHNIATIPYPLED"
ORIGIN
Query Match 100.0%; Score 30; DB 8; Length 3343;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CATGCTTGTTCAACCGTTCGCTGTTGCCA 30
1179 CATGCTTGTTCAACCGTTCGCTGTTGCCA 1208
Db
RESULT 8 158 bp DNA linear PAT 10-SEP-2001
LOCUS AX224396
DEFINITION Sequence 3 from Patent WO0160997.
ACCESSION AX224396
VERSION AX224396.1 GI:15554638
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1
AUTHORS Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
TITLE Male tissue-preferred regulatory region and method of using same
JOURNAL Patent: WO 0160997-A 3 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
source Location/Qualifiers
1..158
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"
ORIGIN
Query Match 80.7%; Score 24.2; DB 6; Length 158;
Best Local Similarity 89.7%; Pred. No. 11;
Matches 26; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CATGCTTGTTCAACCGTTCGCTGTTGCC 29
25 CATGCTTGTTCAACCGTTCGCTGTTGCC 53
Db
RESULT 9 1304 bp DNA linear PAT 24-OCT-2002
LOCUS AX523502
DEFINITION Sequence 90 from Patent WO02064788.
ACCESSION AX523502
VERSION AX523502.1 GI:24412398
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1
AUTHORS Macina,R.A., Recipon,H., Chen,S.Y., Sun,Y. and Liu,C.
TITLE Compositions and methods relating to lung specific genes and
JOURNAL proteins
Patent: WO 02064788-A 90 22-AUG-2002;
Diadexus, Inc. (US)
FEATURES
source Location/Qualifiers
1..1304
/organism="Homo sapiens"

ORIGIN /mol_type="unassigned DNA"
/db_xref="taxon:9606"

Query Match 73.3%; Score 22; DB 6; Length 1304;
Best Local Similarity 83.3%; Pred. No. 97;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CATGCTGTTCAACCGTTCGTTGTTCCA 30
|||||
DB 1155 CTTCTGTTCCCGCCGTTCTGTTCCA 1126

RESULT 10 AC124638 187192 bp DNA linear HTG 05-MAR-2004
AC124638/c MUS MUSCULUS CHROMOSOME 17 CLONE RP23-461H1 map 17, *** SEQUENCING
LOCUS IN PROGRESS ***; 10 unordered pieces.

ACCESSION AC124638 GI:45120281
VERSION HTG; HTGS PHASE1; HTGS FULLTOP; HTGS_ACTIVEFIN.
KEYWORDS MUS MUSCULUS (house mouse)
SOURCE MUS MUSCULUS
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 187192)
AUTHORS Birren, B., Nussbaum, C. and Lander, E.
TITLE Mus musculus chromosome 17, clone RP23-461H1
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 187192)
AUTHORS Birren, B., Linton, L., Nussbaum, C., Lander, E., All, A., Allen, N.,
JOURNAL Anderson, S., Barina, N., Bastien, V., Bloom, T., Boguslavsky, L.,
AUTHORS Boukhalter, B., Brown, A., Camarata, J., Campiano, A., Chang, J.,
AUTHORS Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
AUTHORS Cook, A., Cooke, P., DeRellano, K., Dewar, K., Diaz, J. S., Dodge, S.,
AUTHORS Fero, S., Ferreira, P., Fitzgerald, M., Fitzhugh, W., Gage, D.,
AUTHORS Galagan, J., Gardyna, S., Ginde, S., Gold, S., Goyette, M., Graham, L.,
AUTHORS Grand-Pierre, N., Hago, B., Horton, L., Hulme, W., Iliev, I.,
AUTHORS Johnson, R., Jones, C., Kamat, A., Karatas, A., Kelle, C., LaRocque, K.,
AUTHORS Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K.,
AUTHORS Liu, G., Maclean, C., Macdonald, P., Major, J., Marquis, N.,
AUTHORS Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J.,
AUTHORS Meneses, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,
AUTHORS Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P.,
AUTHORS O'Neill, D., Oliver, J., Peterson, K., Phunhangan, P., Pierre, N.,
AUTHORS Pollara, V., Raymond, C., Retra, R., Rieback, M., Riley, R., Rise, C.,
AUTHORS Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S.,
AUTHORS Schnupack, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N.,
AUTHORS Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S.,
AUTHORS Theodore, J., Topham, K., Travers, M., Travis, N., Triggillo, J.,
AUTHORS Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,
AUTHORS Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (15-JUN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 187192)
REFERENCE 1 (bases 1 to 187192)
AUTHORS Birren, B., Nussbaum, C., Lander, E., Abouelleil, A., Allen, N.,
AUTHORS Anderson, M., Arachchi, H. M., Barina, N., Bastien, V., Bloom, T.,
AUTHORS Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y.,
AUTHORS Collymore, A., Cook, A., Cooke, P., Corum, B., DeRellano, K.,
AUTHORS Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Fero, S.,
AUTHORS Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S.,
AUTHORS Graham, L., Grand-Pierre, N., Hago, B., Horton, L., Hulme, W., Iliev, I.,
AUTHORS Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
AUTHORS Kamat, A., Karatas, A., Kelle, C., Landers, T., Levine, R.,
AUTHORS Lindblad-Toh, K., Liu, X., Liu, A., Mabbitt, R., Maclean, C.,
AUTHORS Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
AUTHORS Meldrum, J., Meneses, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J.,
AUTHORS Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
AUTHORS O'Neill, D., Oliver, J., Peterson, K., Phunhangan, P., Pierre, N.,
AUTHORS Rachupka, A., Ramasamy, U., Raymond, C., Retra, R., Rise, C., Rogov, P.,
AUTHORS Roman, J., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Smith, C.,
AUTHORS Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,

TITILE Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M.,
JOURNAL Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X.,
AUTHORS Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

COMMENT Direct Submission
Submitted (05-MAR-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 5, 2004 this sequence version replaced GI:28893670.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIRB
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: L24115
Center clone name: 461_H_1

* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 9926: contig of 9926 bp in length
* 9927 10026: gap of 100 bp
* 10027 20081: contig of 10055 bp in length
* 20082 20181: gap of 100 bp
* 20182 43359: contig of 23078 bp in length
* 43260 43360: gap of 100 bp
* 43360 45947: contig of 2586 bp in length
* 45948 46047: gap of 100 bp
* 46048 100940: contig of 54893 bp in length
* 100941 101040: gap of 100 bp
* 101040 106719: contig of 5679 bp in length
* 106720 106819: gap of 100 bp
* 106820 140370: contig of 33551 bp in length
* 140371 140470: gap of 100 bp
* 140471 147705: contig of 7235 bp in length
* 147706 147805: gap of 100 bp
* 147806 152565: contig of 4760 bp in length
* 152566 152665: gap of 100 bp
* 152666 187192: contig of 34527 bp in length.

FEATURES

source 1..187192
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="17"
/map="17"
/clone="RP23-461H1"
/clone_lib="RP23-461H1"
Location/Qualifiers

ORIGIN

Query Match 72.0%; Score 21.6; DB 2; Length 187192;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 ATGCTTGTTCACCGTTCGTTGTTCC 29
|||||
DB 52881 ATGCTTGTTCACCGTTCGTTGTTCC 52854

RESULT 11 AC151282 188427 bp DNA linear HTG 24-AUG-2004
AC151282/c MUS MUSCULUS CHROMOSOME 17 CLONE RP23-461H1, *** SEQUENCING IN
LOCUS PROGRESS ***; 12 unordered pieces.
DEFINITION
ACCESSION AC151282
VERSION AC151282.1 GI:51511039

REFERENCE 6 (bases 1 to 129778)
AUTHORS Buell,R.
TITLE Direct Submission
JOURNAL Submitted (09-AUG-2001) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA, rbuell@tigr.org
COMMENT On Jul 11, 2001 this sequence version replaced gi:1404299.
Address all correspondence to:rice@tigr.org

BAC clone OSJNB0093E13 is from *Oryza sativa* chromosome 3
The orientation of the sequence is from SP6 to T7 end of the BAC
clone.
Genes were identified by a combination of several methods: Gene
prediction programs including Fgenesh (<http://www.softberry.com/>),
GENSCAN and Genscan (Chris Buge,
<http://CCR-081.mit.edu/GENSCAN.html>), GenemarkHMM (Mark Borodovsky,
<http://genemark.biology.gatech.edu/Genemark/>), and Genspicr
(Mihela Pertea and Steven Salzberg, contact@tigr.org),
searches of the complete sequence against a peptide database and
the plant EST database at TIGR (<http://www.tigr.org/tdb/tgi.html>).
Annotated genes are named to indicate the level of evidence for
their annotation. Genes with similarity to other proteins are named
after the database hits. Genes without significant peptide
similarity but with EST similarity are named as unknown proteins.
Genes without protein or EST similarity, that are predicted by more
than two gene prediction programs over most of their length are
annotated as hypothetical proteins. Genes encoding tRNAs are
predicted by tRNAscan-SE (Sean Eddy,
<http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats are
identified by RepeatMasker (Arian Smit,
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>).
This BAC overlaps with rice BACs OSJNBa0091J19 (GB:AC084320) and
OSJNB0024J04.

FEATURES
source 1.129778
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/sub_species="japonica"
/db_xref="taxon:39947"
/chromosome="3"
/map="near S13014"
/clone="OSJNB0093E13"
/rpt_family="Gaijin Osz MITE element from gb:D10985 Rice
1701..1821
searize cbp3 gene for carboxypeptidase III (1125 to 1251)
127 nt"
repeat_region complement(2452..2476)
/rpt_family="GAA)n"
4620..5435
/gene="OSJNB0093E13.4"
/note="predicted by fgenesh"
join(<4620..4718,4866..>5435)
/gene="OSJNB0093E13.4"
join(4620..4718,4866..5435)
/gene="OSJNB0093E13.4"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAK72893.1"
/db_xref="GI:14718315"
/translation="MEDVYLSDDSDSVLEHDSVNGSGDSSEGFARSGKLPSSQ
LLTLTKSTQPKRTRRTLPESVPALAGSGHENAELSDVARNNDHVSNAIAADPTF
SGEEDLDDEIDECFDEBDGICENPDEIDFDEELEDDESDLDDEIDECYDVEDK
ICEENPDDEIDFDEEIDCEDEIDDDDDCESLDNDNICFPADIEDIICLETRTKSSST
NIAM"
repeat_region 4640..4775
/rpt_family="(CGG)n"
complement(4730..4835)
/rpt_family="(CGG)n"
11008..11724
/gene="OSJNB0093E13.5"
/note="predicted by fgenesh"
<11008..>11724
mRNA /gene="OSJNB0093E13.5"

CDS 11008..11724
/gene="OSJNB0093E13.5"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAK72895.1"
/db_xref="GI:14718317"
/translation="MEVRRRAVASESGRRRQAYGEQRORGRQLATGRLTALPEA
KGRWSEGERKATVSGRALDGRRRGASGSGRRGRVPLRLPTPAAPVPSV
FGDGRROENDGGNGVYVARSKRLTGMGDLISGFIGMGRFGCGDGLRLRQGW
RAGAYKNAAGAHGGRRESSSGKKEKGVERRGRTMPWGRGREGHERELMHVLDASA
QRWGAAGRHGGLKAGCGC"
repeat_region 12209..12383
/rpt_family="(TGA)n"
complement(12766..16846)
/gene="OSJNB0093E13.7"
/note="highly similar to retrofit gag pol protein
GB:AB82754 GI:2586082 (Oryza longistylata)"
complement(join(<12766..12903,13103..13518,13741..>16846))
/gene="OSJNB0093E13.7"
complement(join(12766..12903,13103..13518,13741..16846))
/gene="OSJNB0093E13.7"
/codon_start=1
/product="putative gag-pol protein"
/protein_id="AAK72882.1"
/db_xref="GI:14718304"
/translation="MASSSSSTAEPANPLFGVQVSEKLTKQNSLMSAQVLTALRGARL
EGHVGTVPEPEALEQEGKEKXTVVPNPARGVGEATDQOVLGFLSLTRIRS
QVAGAPTAAAMKTIENPFSRSHAGATNVALTTQCGOSTVEVYSKRAAGDRI
AATGKPTIDDEELVATITNGLSEEDPVVEALIAKASVTVAEVSQILGFENKVKIRI
ACAATSGRRSGNGSGRGNRGRGTGSGGSGRGGGHORNGRGGTDNRPTQVC
HKGVVADVCMHRVDENVPDEBKGAATHAAYGVDTNMYVOTEADHITGDLKTLTR
EKKGTQDIHPASGEMDIOHIGHSYPTSRPLKNIILVPRKSKMLISVRLVAD
NYAFLEIHQKFLIKDKVTRRTILEGPRSRPLPADDPILKQYVAWVSPFRMGRGL
GASKEPVLRVINQKLLPCSNESPESVCDACCGCKHQLPPKPEFVSNDLEIHS
DVWGPASDVCAKRYYSFIDYKFTWITYLKRSRVEFKFKESQSVBERQFNKIL
GMQDVGSEYQKLSFFKQIGIAHOVSPHHQOQGAVERHRHIVEIGLILLASAM
PLKFWDEAFIATYILINLIPRKVIDPDLPLARLPHQPDYKSLRFGCAQCPNLPNP
THKLOFRSKQCVLFGFSNLHKFKCLVATQVVISRPVDEDEFPFSKLOPNAGAR
LRAEISLPLADVTSHASDQGVANSDHMHNIILPVYDSSFNEDAADDTDSDVAE1
NABOPDGIHPASTGVHTAEVEVAASDSSPDORRVAQAPQATVTEPAAATSAVO
SSRASRSIASHQIGDTPSPGPHHDAAGSSAAPEVATATGCGIDINARPKTRLQSGIR
KKIYITDGTIVYGCPTSSGEPONLAELAGDNWKEAMKEVYALMKSTHMLVPRKG
INIIDCKWVYIKRKAQSLDRYKGRILAKAPKORYGIDVDEPSPVKAATKIDVLS
LVASQDTSLPFYDKKGVTFMILVYVDIIVSSSEKATYALDLOLQCFALKIDGLH
PFLGVEYKVDGAVLIDQEKYASDVLRRVSGDCKPVNTPISIKSKSOGGALQGN
DSTQRSIVGALQVLTILRPDISFSVKNQPTLYLGVHKNKQPLDQVLTAKNTRHMQW
QOKLCGYKHYFRNLY"
repeat_region complement(113497..13708)
/rpt_family="Rice retrotransposon Tos7 DNA for reverse
transcriptase, partial cds gi|1621482|dbj|D85871.1|D85871"
complement(16049..16146)
repeat_region /rpt_family="(CGG)n"
17933..19954
/gene="OSJNB0093E13.6"
/note="predicted by fgenesh"
<17933..>19954
mRNA /gene="OSJNB0093E13.6"
17933..19954
/gene="OSJNB0093E13.6"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAK72884.1"
/db_xref="GI:14718306"
/translation="MEDVYLSDDSDSVLEHDSVNGSGDSSEGFARSGKLPSSQ
LLTLTKSTQPKRTRRTLPESVPALAGSGHENAELSDVARNNDHVSNAIAADPTF
SGEEDLDDEIDECFDEBDGICENPDEIDFDEELEDDESDLDDEIDECYDVEDK
ICEENPDDEIDFDEEIDCEDEIDDDDDCESLDNDNICFPADIEDIICLETRTKSSST
NIAM"
repeat_region 4640..4775
/rpt_family="(CGG)n"
complement(4730..4835)
/rpt_family="(CGG)n"
11008..11724
/gene="OSJNB0093E13.5"
/note="predicted by fgenesh"
<11008..>11724
mRNA /gene="OSJNB0093E13.5"

MOOAAARNAAAAVRADAIDSVMAGGLAODVEALSGATGRLMAAAPSPA
 VGVDSASNTAAAAFLPDTIALQDGVSLAVFILLRFVVTLLIRICALN"
 repeat_region
 18749..18856
 /rpt_family="(GGA)n"
 repeat_region
 complement(20115..20185)
 /rpt_family="AT_rich"
 21519..21947
 /gene="OSJNB0093E13.8"
 /note="Predicted by fgenesh"
 mRNA
 <21519..21947
 /gene="OSJNB0093E13.8"
 21519..21947
 /gene="OSJNB0093E13.8"
 /codon_start=1
 /product="hypothetical protein"

Query Match 70.7%; Score 21.2; DB 8; Length 129778;
 Best Local Similarity 88.5%; Pred. No. 2.2e+02;
 Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 CTTGTTCAACCGTTCGTTGTTCCA 30
 Db 20209 CATGTTAACCGTTCGTTGTTCCA 20184

RESULT 13
 AP006548 140729 bp DNA linear PLN 12-JUL-2003
 LOCUS Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 9,
 DEFINITION PAC clone: P0705E11.
 ACCESSION AP006548
 VERSION AP006548.1 GI:32490470
 KEYWORDS

SOURCE
 ORGANISM
 Oryza sativa (japonica cultivar-group)
 Oryza sativa (japonica cultivar-group)
 Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Erihartoideae; Oryzaceae; Oryza.

REFERENCE
 AUTHORS
 Oliveira,A.C., Matos,L.T., Carvalho,F.F., Shimano,A., Zimmer,P.D.,
 Malone,G. and Dellagostin,O.
 Oryza sativa n1pdonbare genomic DNA, chromosome 9, PAC
 clone:P0705E11
 Published only in Database (2003)
 2 (bases 1 to 140729)
 Oliveira,A.C., Matos,L.T., Carvalho,F.F., Shimano,A., Zimmer,P.D.,
 Malone,G. and Dellagostin,O.
 Direct Submission

TITLE
 JOURNAL
 Submitted (08-JUL-2003) Antonio C Oliveira, Universidade Federal de
 Pelotas, Centro de Genomica e Fitomelhoramento, Departament
 Fitotecnia; Campus Universitario do Capao do Leao, Pelotas,
 Pelotas, Rio Grande do Sul 96010900, Brasil
 (E-mail:acostol@terra.com.br,
 URL:http://www.ufpel.tche.br/faem/fitotecnia/fitomelhoramento/
 Tel:55-53-275-7263; Fax:55-53-275-9031)
 Gene were predicted from the integrated results of the following:
 GENSCAN (http://CCR-081.mt.edu/GENSCAN.html), FGENESH
 (http://www.softberry.com/), Genemark.hmm
 (http://opal.biology.gatech.edu/Genemark/), Glimmer
 (http://www.tigr.org/cdb/glimmer/glmrform.html), RicheHM
 (http://xgp.dna.affrc.go.jp/RicheHM/), SplinePredictor
 (http://bioinformatics.lastate.edu/cgi-bin/sp.cgi), BLASTN and
 BLASTX. The genomic sequence was searched against NCBI NonRedundant
 Protein database, nr

COMMENT
 (ftp://ncbi.nlm.nih.gov/blast/db) and the EST sequence database at
 RGP. Protein homologies of the coding regions were searched against
 NCBI NonRedundant Protein database with BLASTP. ESTs represent the
 identified cDNA sequences using BLASTN with the corresponding DBS
 accession no. and RGP clone ID.
 A gene with identity or significant homology to a protein is
 classified based on the protein name to indicate the homology level
 such as same name, 'putative-' and '-like protein'. A gene without
 significant homology to any protein but with EST homology (covering

FEATURES

source
 almost the entire length of partial sequence) is classified as an
 'unknown' protein. A gene predicted by two or more gene prediction
 programs is classified as a 'hypothetical' protein according to
 IRGSP standard.
 The orientation of the sequence is from SP6 to T7 of the PAC clone.
 This sequence of P0705E11 clone has an overlap with OSJNB0069D16
 clone (DDBJ: AP005882) at 5' end and an overlap with OSJNB0069D16
 OSJNBa0038K02 (DDBJ: AP005862) at 3' end.
 Location/Qualifiers

1..140729
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="genomic DNA"
 /cultivar="n1pdonbare"
 /db_xref="taxon:39947"
 /chromosome="9"
 /clone="P0705E11"
 complement(1590..2675)
 /note="3'LTR"
 3382..4071
 /gene="P0705E11.1"
 3382..4071
 /gene="P0705E11.1"
 /note="probably inactive due to 3' exon missing in CDS
 probably inactive due to 5' exon missing in CDS
 pseudogene, orf of gypsy-type retrotransposon RIRE2"
 /pseudo
 complement(7148..9310)
 /gene="P0705E11.2"
 complement(7148..9310)
 /gene="P0705E11.2"
 /note="probably inactive due to 5' exon missing in CDS
 pseudogene, GAG-POL precursor of gypsy-type
 retrotransposon RIRE2"
 /pseudo
 complement(13300..14385)
 /note="5'LTR"
 20095..21408
 /gene="P0705E11.3"
 <20095..21408
 /gene="P0705E11.3"
 /note="start and end point are not identified"
 20095..21408
 /gene="P0705E11.3"
 /codon_start=1
 /product="putative anthranilate N-benzoyltransferase"
 /protein_id="BAC79154.1"
 /db_xref="GI:32490471"
 /translation="MEVKTLSSKLVKPAVNGVAAAPVETPLISFDKVTYKQMAI
 IYAPPPASTAIIEKGLAVIAQYRAFQGLGESPDSGAAYVINDRGARLBAAYDA
 DIVDAPAPKPTBELRLHPDLLEGLOEVVLQLTRCSGLAVGFTSNVADGHAITS
 NFLVWAGPRTKGLPMGAPVHNHAALEFKRPSPHEDHNRNEEYLPAAAGDSHGSD
 GCAADIVIHKAHFTKDFIAGLAAASBERGAPFSFETILAHLMWTRAGSLSPDE
 ASTIRLSVGRIRLAPAEYFGNVLVYAPRAVTVGDLTRPLKHAQVHDAVARVD
 AYPRSLDRLALSGAGDKEGLAPSAVLKDVLCPMAYVDSWLPFPPELDKFGCSPTTF
 MGSYPTBEMLTVLPSYILGDSVDAFVFNHLEFKECTSME"
 33065..34423
 /gene="P0705E11.4"
 <33065..34423
 /note="start and end point are not identified"
 33065..34423
 /gene="P0705E11.4"
 /codon_start=1
 /product="putative anthranilate N-benzoyltransferase"
 /protein_id="BAC79155.1"
 /db_xref="GI:32490472"
 /translation="MEVKTLSSRLRPSVPASAAAPBEFVSSMFDKTYTQMOMAI
 IYAPPPGSAVADIERGLAVLGVRPLFAGQVVRGGGELRGVNLNDHGRALVEACVDS
 SLADIAPAPSPVIVRLHPDSLEGEIEVQVQVLTFRFAGCSGLAVGFTSNVADGHAITS
 DFLVWAGPRAAGLAAVATAPAPPHHPGFRPRDPLVEFBERGVETYPAPPAAGVD
 GVGGDHKQOHGGEASHGIVIKAHPTKQFIRLRLAAASBERGAPFSFETILAHLMWTRAGSLSPDE
 WRTWTRAGSLNPSITIRISVGRIRLSAPGVNGLVMAFPRAVTVGDLGRPL
 KLAQVITHDAVARADAAVFRSVFVPAASGAVGBGLAPAVLKDVLCPLDLEVDSWLTF

Gene 42989. 445252
/gene="P0705E11.5"
misc_feature 42989. 445252
/gene="P0705E11.5"
/note="Anticirrhium majus shows structural homology to the maize transposon En/Spm and has no sequence specificity of insertion
probably inactive due to 3' exon missing in CDS
probably inactive due to 5' exon missing in CDS
probably inactive due to including frameshift(s) in CDS
pseudogene, transposable element Tami"
/pseduo
Gene 49006. 52703
/gene="P0705E11.6"
MRNA join(49006, 50427, 50502, .51258, 51322, .51598, 51708, .51981, 51978, .52088, 52173, .52245, 52340, .52438, 52518, .525703)
/gene="P0705E11.6"
/note="start and end point are not identified"
CDS join(49006, 50427, 50502, .51258, 51322, .51598, 51708, .51981, 51978, 52088, 52173, .52245, 52340, .52438, 52518, .52703)
/gene="P0705E11.6"
/codon_start=1
/product="reverse transcriptase - like protein"
/protein_id="BAC79156.1"
/translation="MADRDKEOILYDTIAGSSGYWMEERENGNDPOYINBEGNVERD
AEGNOQGHVEDVEGNOBERASGSGPVGQKRRAGQRAAKKEGHIITTEBEDRP
SAPAAANVYRHSGVAVRDVNPVSTVYMRTPARGHSEFVDSKEMLTMTMETE
TLPTGEETVRRWTLKQMAEQPFQKGLDYQKYLKQETNPFLPKLRHNDQFVAY
KTGEGKAMERKENAKKKYKHHGLSGGYSVAMPQEMEASLERGEREPATNAP
ERSKGVYAHGNTNPADGSLVFGYIOEARRLTDVEASOGTFRPDRBELTALA
LOTPEHPERTGKGVIPKIKGKEDITHVSRMSKSDTAKIADIAPRYSYELMO
EVAUKDQEMAAHSHDPOPTTPRATVPSGRSKASGYGQSDMAPMOTPDSTO
CPVDITRTFCEIHIPKULSTVASKMALPDPSTYTCRTIPAGYSKVELYELVEG
AYEDLELPGDGETHLRDTCHAILMRKRYTILPEKQASAPSPALPSPDPA
PSPAPAPAPPPQAPASTPDQAPITPRPAPPPAPAPAPAPAPAPAPAPAPAPAP
ATKAKVAAKANDPDGYCTOEBELDAVASVKGQFPRSEKRTIDISVRFKFM
SASVEALDKIDERTLKAKSGSKAVPQGEOPNOEIPLTGVAPEKKEVKMY
ELGRVYBELLOSILPTQMYKFOHLMVAMRBTGEMGARIRDTDLQDDILMINR
GIVELYOALDVCIMSCWITIMEIOPARRRVPRTGSDIPMKVNMALDOYPOETDN
LVHLKAKHYTFTLLPNTPEFHVALLIDIEACTAVNTDSMKRKSSTFDKVELDR
AMVRRHLVVRKMERLRKRFPCQAOKQGMNLGCVYVEYCHADQIITRELD
IMRMRNLTHREIAAVOBOLMPFINEIILDPKEFYDNTIHRSLASLAATTS
KS"
Gene 57466. 64290
/gene="P0705E11.7"
MRNA join(57466, 57935, 58078, .58192, 59352, .59443, 59596, .59709, 59896, .59950, 60105, .60215, 60983, .61036, 61119, .61294, 61380, .61487, 62089, .62162, 62257, .62432, 62926, .63021, 63291, .63371, 63470, .63523, 64003, .64137, 64210, .64290)
/gene="P0705E11.7"
/note="start and end point are not identified"
CDS join(57466, 57935, 58078, .58192, 59352, .59443, 59596, .59709, 59896, .59950, 60105, .60215, 60983, .61036, 61119, .61294, 61380, .61487, 62089, .62162, 62257, .62432, 62926, .63021, 63291, .63371, 63470, .63523, 64003, .64137, 64210, .64290)
/gene="P0705E11.7"
/note="contains EST(s) : AU101437(E31530), AU101438(E31530)"
/codon_start=1
/product="putative serine/chreonine-protein kinase ctrl"
/protein_id="BAC79157.1"
/db_xref="GI:32490474"
/translation="MAVEEAASVCGSHAAASAAAGRCGGCGGCGATSSSSASAGAA
AAAAAKQOOORHKLVEYTVELRLRLDSGVPEAKRGFDELMHNFNLRPAVADY
NVERVEDYTRRLLEOKDAORAPAAVAGVYSLDNDONDAQSNRGAETSRVLSRL
LNROSHHPRAFGSSNTNLEALAEAKSGQSDHSDSNVNRPMHETFTSTIDPK
LISELTSLGELGANTIOBAHAFSTNDYSIDVFPVYGMHDEETEDILBSYKRIKID
ETQGSSTHSHSVESNMQIGENSAOHVEIPDGAASWETIDVLKLFKNGASGSG
DLRYGTSQOVAIKVLKPERINADMQGVFLPALLGVNDVSGMSYLIQNNIIRDTKAIL
CIVETYSGGGVYDLHKHKGVFLPALLGVNDVSGMSYLIQNNIIRDTKAIL
MDENGVYVAAPGVARVYAGSGVFLPALLGVNDVSGMSYLIQNNIIRDTKAIL

```

gene
    ID: E1LTKGI1PVEYVLTPLQANAVONVGKAPRTPKNAHAKLSLILCKMCGQEAPRHPDSE
    ILETLTORIAAEVDEBDEHGKKKEIKELGSLPSALRGRC"H"
    65979..84215
    /gene="P0705E11.8"
mRNA
    join(65979..66134,66228..66279,67599..75119,75210..75
    78937..76241,76395..76475,76728..77036,78187..78573,
    78721..79524,79613..79975,80194..80424,80691..81035,
    81251..81469,81516..81572,81735..82079,82255..82513,
    82513..82513)
Query Match      70.7%; Score 21.2; DB 8; Length 140729;
Best Local Similarity 88.5%; Pred. No. 2.2e+02;
Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy      5 CTTGTTCACCGTTCGCTTGTTCCA 30
        |||||
Db       55011 CATGTTAACGGTTCGCTTAATCCCA 55036

RESULT 14
AP005710      153675 bp   DNA      linear   HTG 06-SEP-2002
LOCUS         Oryza sativa (japonica cultivar-group) chromosome 9 clone P0646B07,
DEFINITION    ** SEQUENCING IN PROGRESS **.
ACCESSION     AP005710
VERSION       AP005710.1 GI:22775481
KEYWORDS      HTG; HTGS PHASE2.
SOURCE        Oryza sativa (japonica cultivar-group)
ORGANISM      Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyte;
               Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
               Ehrhartoidae; Oryzaceae; Oryza.
REFERENCE     1
AUTHORS       Sasaki,T., Matsumoto,T. and Katayose,Y.
TITLE         Oryza sativa nipponbare(GAS) genomic DNA, chromosome 9, PAC
              clone:P0646B07
JOURNAL       Published Only in Database (2002)
REFERENCE     2 (bases 1 to 153675)
AUTHORS       Sasaki,T., Matsumoto,T. and Katayose,Y.
TITLE         Direct Submission
JOURNAL       Submitted (05-SEP-2002) Takuji Sasaki, National Institute of
              Agricultural Sciences, Rice Genome Research Program, Kannondai
              2-1-2, Tsukuba, Ibaraki,305-8602, Japan
              (E-mail:sasaki@nias.affrc.go.jp, URL:http://rpg.dna.affrc.go.jp/,
              TEL:81-298-38-7441, Fax:81-298-38-7468)
COMMENT       NOTE: It currently consists of 1 contigs. Gaps between the contigs
              to be correct as given, however the sizes of the gaps between them
              are represented as runs of N. The order of the pieces is believed
              sequence will be replaced by the finished sequence as soon as it is
              available and the accession number will be preserved.
              * NOTE: This is a 'working draft' sequence.
              * This sequence will be replaced
              * by the finished sequence as soon as it is available and
              * the accession number will be preserved.
              Location/Qualifiers
                source          1..153675
                                /organism="Oryza sativa (japonica cultivar-group)"
                                /mol_type="genomic DNA"
                                /cultivar="Nipponbare"
                                /db_xref="taxon:39947"
                                /chromosome="9"
                                /clone="P0646B07"

FEATURES
    SOURCE
        1..153675
            /organism="Oryza sativa (japonica cultivar-group)"
            /mol_type="genomic DNA"
            /cultivar="Nipponbare"
            /db_xref="taxon:39947"
            /chromosome="9"
            /clone="P0646B07"

ORIGIN
Query Match      70.7%; Score 21.2; DB 2; Length 153675;
Best Local Similarity 88.5%; Pred. No. 2.2e+02;
Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy      5 CTTGTTCACCGTTCGCTTGTTCCA 30
        |||||
Db       19380 CATGTTAACGGTTCGCTTAATCCCA 19405
```



```
misc_feature
18062..18190,18223..18322,18729..18861,18967..19111)
/gene="OJ1148_D05.5"
/complement(join(15074..15178,17827..17877,17931..18029,
18062..18190,18223..18322,18729..18861,18967..19111))
/gene="OJ1148_D05.5"
/note="hypothetical ORF
predicted by GENSCAN
this category is not included in IRGSP standard"
complement(25292..25586)
/gene="OJ1148_D05.6"
/complement(join(25292..25412,25471..25586))
/gene="OJ1148_D05.6"
/note="start and end point are not identified"
complement(join(25292..25412,25471..25586))
/gene="OJ1148_D05.6"
/note="predicted by GeneMark.hmm etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD07688.1"
/db_xref="GI:41053262"
/translation="MAHSGAGLKAATVTRGDAMARQRRTGREALMRRRRRDCRRV
DGAATATRHDAARQRQSEATATAPVRAGLENG"
/complement(join(26131..26319,28092..28139))
/gene="OJ1148_D05.7"
/complement(join(26131..26319,28092..28139))
/gene="OJ1148_D05.7"
/note="hypothetical ORF
predicted by GENSCAN
this category is not included in IRGSP standard"
30668..30991
/gene="OJ1148_D05.8"
<30668..30991
/gene="OJ1148_D05.8"
/note="start and end point are not identified"
30668..30991
/gene="OJ1148_D05.8"
/note="predicted by FGENESH etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD07689.1"
/db_xref="GI:41053263"
/translation="MPCISPSCKATTCRTPGNHHGTSLELAASRRASLWEENVLP
YNFPPLSPSYFSSSCOHILAVMGVEVANVTKVVGEGRKIKREEDEEWLNGSHA
ILDPT"
32784..35862
/gene="OJ1148_D05.9"
/join(32784..33234,33666..34250,35333..35862)
/gene="OJ1148_D05.9"
/note="supported by full-length cDNA(s) : AK060695"
/join(32872..33234,33666..34250,35333..35506)
/gene="OJ1148_D05.9"
/note="contains full-length cDNA(s) : AK060695"
/codon_start=1
/product="putative basic-helix-loop-helix transcription
factor"
/protein_id="BAD07690.1"

gene
32784..35862
mRNA
/join(32784..33234,33666..34250,35333..35862)
/gene="OJ1148_D05.9"
/note="supported by full-length cDNA(s) : AK060695"
/join(32872..33234,33666..34250,35333..35506)
/gene="OJ1148_D05.9"
/note="contains full-length cDNA(s) : AK060695"
/codon_start=1
/product="putative basic-helix-loop-helix transcription
factor"
/protein_id="BAD07690.1"

CDS
32784..35862
gene
32784..35862
mRNA
/join(32784..33234,33666..34250,35333..35862)
/gene="OJ1148_D05.9"
/note="supported by full-length cDNA(s) : AK060695"
/join(32872..33234,33666..34250,35333..35506)
/gene="OJ1148_D05.9"
/note="contains full-length cDNA(s) : AK060695"
/codon_start=1
/product="putative basic-helix-loop-helix transcription
factor"
/protein_id="BAD07690.1"

Query Match 70.7%; Score 21.2; DB 8; Length 155146;
Best Local Similarity 88.5%; Pred. No. 2.2e+02;
Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

Search completed: September 16, 2005, 03:01:24
Job time : 105.667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using bw model

Run on: September 15, 2005, 08:10:17 ; Search time 14.5242 Seconds
(without alignments)
12227.321 Million cell updates/sec

Title: US-10-713-381-1_COPY_1179_1208

Perfect score: 30
Sequence: 1 catgctgtgtcaacgctcgtctgtcca 30

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	100.0	30	5 AAH76335	Aah76335 Z. mays M
2	30	100.0	255	5 AAH76340	Aah76340 Z. mays M
3	30	100.0	1394	2 AAH07408	Aax07408 Zea mays
4	30	100.0	1394	2 AAH07409	Aax07409 Zea mays
5	30	100.0	1394	5 AAH76332	Aah76332 Z. mays M
6	30	100.0	1394	5 AAH76333	Aah76333 Z. mays M
7	24.2	80.7	158	5 AAH76334	Aah76334 Z. mays M
8	22	73.3	1304	6 AB075351	Abq75351 Human lun
9	20.4	68.0	1542	6 ABZ13819	Abz13819 Arabidops
10	20.2	67.3	2397	8 ACA44013	Ac444013 Prokaryot
11	20	66.7	2000	12 ADJ40785	Adj40785 Plant CDN
12	19.6	65.3	2000	8 ADA72061	Ada72061 Rice gene
13	19.6	65.3	2000	8 ADA71994	Ada71994 Rice gene
14	19.6	65.3	2000	8 ADA71736	Ada71736 Rice gene
15	19.6	65.3	2000	8 ADA72530	Ada72530 Rice gene
16	19.6	65.3	2000	12 ADJ40977	Adj40977 Plant CDN
17	19.4	64.7	2240	2 AAQ47668	Aaq47668 Sequence
18	19.4	64.7	2243	4 AAF99956	Aaf99956 Rat brain
19	19.4	64.7	2243	9 ACD26214	Ac26214 CDNA enco
20	19.4	64.7	349980	6 ABQ81845	Abq81845 Bifidobac

C	21	19.2	64.0	2000	8 ADA73340	Ada73340 Rice gene
C	22	19	63.3	1324	8 ADA71777	Ada71777 Rice gene
C	23	19	63.3	1388	3 AAC44356	Aac44356 Arabidops
C	24	19	63.3	1821	10 ADE93879	Ade93879 Ixodes sc
C	25	19	63.3	5671	4 AAK90190	Aak90190 Human dig
C	26	19	63.3	5671	5 ABA20915	Ab20915 Human dir
C	27	19	63.3	5671	5 AAS39832	Aas39832 Genomic s
C	28	19	63.3	5671	9 ADB32792	Adb32792 Human nov
C	29	19	63.3	198522	11 ACN44010	Acn44010 Human gen
C	30	18.8	62.7	883	6 ABR65288	AbR65288 Arabidops
C	31	18.8	62.7	883	10 ADD30938	Add30938 Plant yie
C	32	18.8	62.7	883	10 ADE37244	Ade37244 Plant yie
C	33	18.8	62.7	883	12 ADI41918	Adi41918 Plant tra
C	34	18.8	62.7	1083	4 AAS53644	Aas53644 Helicobac
C	35	18.8	62.7	1083	8 ACA34769	Ac34769 Prokaryot
C	36	18.8	62.7	1899	10 ABX06821	Abx06821 S. pneumo
C	37	18.8	62.7	1902	4 AAS55583	Aas55583 Streptoco
C	38	18.8	62.7	1902	8 ACA49918	Ac49918 Prokaryot
C	39	18.8	62.7	1908	12 ADJ63835	Adj63835 Plant 11p
C	40	18.8	62.7	2000	12 ADJ41535	Adj41535 Plant CDN
C	41	18.8	62.7	2247	13 ADS57809	Ad57809 Bacterial
C	42	18.8	62.7	2920	11 ADM02356	Adm02356 Human CDN
C	43	18.8	62.7	3339	10 ACF79841	AcF79841 Mouse ani
C	44	18.8	62.7	7812	3 AAZ61238	Aaz61238 DNA encod
C	45	18.8	62.7	9541	2 AAV52265	Aav52265 Streptoco

ALIGNMENTS

RESULT 1	AAH76335	standard; DNA, 30 BP.
ID	AAH76335	
AC	AAH76335;	
XX		
DT	29-OCT-2001	(first entry)
XX		
DE	Z. mays Ms45 male tissue-preferred regulatory region fragment.	
XX		
KW	Ms45; male tissue; regulatory region; transcription; male fertility;	
KM	hybrid seed; ds.	
XX		
OS	Zea mays.	
XX		
PN	W0200160997-A2.	
XX		
PD	23-AUG-2001.	
XX		
PF	13-FEB-2001; 2001WO-US004527.	
XX		
PR	15-FEB-2000; 2000US-00504487.	
XX		
PA	(PION-) PIONEER HI-BRED INT INC.	
XX		
PI	Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;	
XX		
DR	WPI; 2001-514772/56.	
XX		
PT	A male tissue-preferred regulatory region comprising nucleotide sequences	
PT	essential for initiating transcription of the Ms45 gene useful for	
PT	mediating fertility in a male plant.	
XX		
PS	Claim 5; Page 47; 50pp; English.	
XX		
CC	The invention provides a male tissue-preferred regulatory region (I)	
CC	comprising nucleotide sequences essential for initiating transcription of	
CC	the Ms45 gene. A method of mediating male fertility in a plant is	
CC	provided that involves introducing an expression vector comprising a	
CC	promoter operably linked to (I) into a plant where the exogenous gene	
CC	impacts male fertility of the plant and (II) controls expression of the	
CC	exogenous gene. A method of producing hybrid seeds is also provided. The	
CC	present sequence represents a DNA fragment -152 to -161 bases upstream of	

CC the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
 CC nucleotide sequence
 XX
 SQ Sequence 30 BP; 4 A; 9 C; 5 G; 12 T; 0 U; 0 Other;

Query Match 100.0%; Score 30; DB 5; Length 30;
 Best Local Similarity 100.0%; Pred. No. 0.004;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
 |||||
 DB 1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30

RESULT 2

AAH76340
 ID AAH76340 standard; DNA; 255 BP.

AC AAH76340;

XX 29-OCT-2001 (first entry)

DE Z. mays Ms45 promoter fragment.

XX Ms45; male tissue; regulatory region; transcription; male fertility;

KM hybrid seed; promoter; de.

XX Zea mays.

OS WO200160997-A2.

XX 23-AUG-2001.

PD 13-FEB-2001; 2001WO-US004527.

XX 15-FEB-2000; 2000US-00504487.

XX (PION-) PIONEER HI-BRED INT INC.

XX PI Albertsen MC, Fox TW, Garmaat CW, Huffman G, Kendall TL;

XX WPI; 2001-514772/56.

XX A male tissue-preferred regulatory region comprising nucleotide sequences
 PT essential for initiating transcription of the Ms45 gene useful for
 PT mediating fertility in a male plant.

XX Example 5; Fig 8; 50pp; English.

XX The invention provides a male tissue-preferred regulatory region (I)
 CC comprising nucleotide sequences essential for initiating transcription of
 CC the Ms45 gene. A method of mediating male fertility in a plant is
 CC provided that involves introducing an expression vector comprising a
 CC promoter operably linked to (I) into a plant where the exogenous gene
 CC impacts male fertility of the plant and (I) controls expression of the
 CC exogenous gene. A method of producing hybrid seeds is also provided. The
 CC present sequence represents a Z. mays Ms45 promoter fragment

XX SQ Sequence 255 BP; 59 A; 86 C; 39 G; 71 T; 0 U; 0 Other;

Query Match 100.0%; Score 30; DB 5; Length 255;
 Best Local Similarity 100.0%; Pred. No. 0.0051;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
 |||||
 DB 39 CATGCTTGTTCAACCGTTCGTTGTTCCA 68

RESULT 3

AAK07408
 ID AAK07408 standard; DNA; 1394 BP.

AC AAK07408;

XX 08-JUN-1999 (first entry)

XX Zea mays Ms45 male tissue-preferred regulatory region.

XX Ms45; male; tissue-preferred; regulatory region; plant cells;

KM plant tissue; differentiated; maize; hybrid seed; fertility; ss.

XX Zea mays.

OS WO9859061-A1.

XX 30-DEC-1998.

XX 19-JUN-1998; 98WO-US012895.

XX 23-JUN-1997; 97US-00880499.

XX (PION-) PIONEER HI-BRED INT INC.

XX PI Albertsen MC, Fox TW, Garmaat CW, Huffman GA, Kendall TL;

XX WPI; 1999-105628/09.

XX New nucleic acid encoding a Ms45 male tissue-preferred regulatory region
 PT - useful in mediating plant fertility, especially hybrid seed production.

XX Claim 2; Page 22-23; 39pp; English.

XX The sequence is that encoding an Ms45 male tissue-preferred regulatory
 CC region. It may be used in the construction of a vector for a method of
 CC producing exogenous genes in a male tissue- preferred manner, which is
 CC useful in restoring or conferring fertility, such as in hybrid seed
 CC production. In conferring fertility, a monocot/dicot plant is transformed
 CC with the exogenous nucleotide sequence (a male sterility gene, preferably
 CC Ms45), which encodes a product selected from auxins, rolb and dipteria
 CC toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
 CC and infertile plants

XX SQ Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;

Query Match 100.0%; Score 30; DB 2; Length 1394;
 Best Local Similarity 100.0%; Pred. No. 0.0063;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
 |||||
 DB 1179 CATGCTTGTTCAACCGTTCGTTGTTCCA 1208

RESULT 4

AAK07409
 ID AAK07409 standard; DNA; 1394 BP.

XX AAK07409;

XX 08-JUN-1999 (first entry)

XX Zea mays Ms45 male tissue-preferred regulatory region.

XX Ms45; male; tissue-preferred; regulatory region; plant cells;

KM plant tissue; differentiated; hybrid seed; fertility; ss.

XX Zea mays.

OS WO9859061-A1.

XX 30-DEC-1998.

XX 19-JUN-1998; 98WO-US012895.

XX 23-JUN-1997; 97US-00880499.

```
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Albertsen MC, Fox TW, Garnaat CW, Huffman GA, Kendall TL;
XX WPI; 1999-105628/09.
XX
XX New nucleic acid encoding a Ms45 male tissue-preferred regulatory region
XX - useful in mediating plant fertility, especially hybrid seed production.
XX
XX Claim 3; Page 23-24; 39pp; English.
XX
XX The sequence is that encoding an Ms45 male tissue-preferred regulatory
XX region. It may be used in the construction of a vector for a method of
XX producing exogenous genes in a male tissue- preferred manner, which is
XX useful in restoring or conferring fertility, such as in hybrid seed
XX production. In conferring fertility, a monocot/dicot plant is transformed
XX with the exogenous nucleotide sequence (a male sterility gene, preferably
XX Ms45), which encodes a product selected from auxins, rolb and diphenolia
XX toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
XX and infertile plants
XX
XX Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 30; DB 2; Length 1394;
XX Best Local Similarity 100.0%; Pred. No. 0.0063;
XX Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
XX 1179 CATGCTTGTTCAACCGTTCGTTGTTCCA 1208
XX
XX
XX RESULT 5
XX AAH76332
XX ID AAH76332 standard; DNA; 1394 BP.
XX
XX AAH76332;
XX
XX 29-OCT-2001 (first entry)
XX
XX Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.
XX
XX Ms45; male tissue; regulatory region; transcription; male fertility;
XX hybrid seed; de.
XX
XX Zea mays.
XX
XX WO200160997-A2.
XX
XX 23-AUG-2001.
XX
XX 13-FEB-2001; 2001MO-US004527.
XX
XX 15-FEB-2000; 2000US-00504487.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX WPI; 2001-514772/56.
XX
XX A male tissue-preferred regulatory region comprising nucleotide sequences
XX essential for initiating transcription of the Ms45 gene useful for
XX mediating fertility in a male plant.
XX
XX Claim 4; Page 46; 50pp; English.
XX
XX The invention provides a male tissue-preferred regulatory region (I)
XX comprising nucleotide sequences essential for initiating transcription of
XX the Ms45 gene. A method of mediating male fertility in a plant is
XX provided that involves introducing an expression vector comprising a
XX promoter operably linked to (I) into a plant where the exogenous gene
XX impacts male fertility of the plant and (I) controls expression of the
XX exogenous gene. A method of producing hybrid seeds is also provided. The
XX present sequence represents a nucleic acid sequence encoding an Ms45 male
XX tissue preferred regulatory region from Z. mays
XX
XX Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 30; DB 2; Length 1394;
XX Best Local Similarity 100.0%; Pred. No. 0.0063;
XX Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
XX 1179 CATGCTTGTTCAACCGTTCGTTGTTCCA 1208
XX
XX
XX RESULT 7
```

```
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a nucleic acid sequence encoding an Ms45 male
CC tissue preferred regulatory region from Z. mays
XX
XX Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 30; DB 5; Length 1394;
XX Best Local Similarity 100.0%; Pred. No. 0.0063;
XX Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
XX 1179 CATGCTTGTTCAACCGTTCGTTGTTCCA 1208
XX
XX
XX RESULT 6
XX AAH76333
XX ID AAH76333 standard; DNA; 1394 BP.
XX
XX AAH76333;
XX
XX 29-OCT-2001 (first entry)
XX
XX Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.
XX
XX Ms45; male tissue; regulatory region; transcription; male fertility;
XX hybrid seed; de.
XX
XX Zea mays.
XX
XX WO200160997-A2.
XX
XX 23-AUG-2001.
XX
XX 13-FEB-2001; 2001MO-US004527.
XX
XX 15-FEB-2000; 2000US-00504487.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX WPI; 2001-514772/56.
XX
XX A male tissue-preferred regulatory region comprising nucleotide sequences
XX essential for initiating transcription of the Ms45 gene useful for
XX mediating fertility in a male plant.
XX
XX Claim 4; Page 47; 50pp; English.
XX
XX The invention provides a male tissue-preferred regulatory region (I)
XX comprising nucleotide sequences essential for initiating transcription of
XX the Ms45 gene. A method of mediating male fertility in a plant is
XX provided that involves introducing an expression vector comprising a
XX promoter operably linked to (I) into a plant where the exogenous gene
XX impacts male fertility of the plant and (I) controls expression of the
XX exogenous gene. A method of producing hybrid seeds is also provided. The
XX present sequence represents a nucleic acid sequence encoding an Ms45 male
XX tissue preferred regulatory region from Z. mays
XX
XX Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 30; DB 5; Length 1394;
XX Best Local Similarity 100.0%; Pred. No. 0.0063;
XX Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
XX 1179 CATGCTTGTTCAACCGTTCGTTGTTCCA 1208
XX
XX
XX RESULT 7
```

```

AAH76334
ID AAH76334 standard; DNA; 158 BP.
XX
XX
AC AAH76334;
XX
XX
DT 29-OCT-2001 (first entry)
XX
XX
DE Z. mays Ms45 male tissue-preferred regulatory region fragment.
XX
XX
KM Ms45; male tissue; regulatory region; transcription; male fertility;
XX
XX
KM hybrid seed; ds.
XX
XX
OS Zea mays.
XX
XX
PN WO200160997-A2.
XX
XX
PD 23-AUG-2001.
XX
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
XX
PR 15-FEB-2000; 2000US-00504487.
XX
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
XX
DR WPI; 2001-514772/56.
XX
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the MS45 gene useful for
PT mediating fertility in a male plant.
XX
XX
PS Claim 5; Page 47; 50pp; English.
XX
XX
CC The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the MS45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a DNA fragment -38 to -195 bases upstream of
CC the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
CC nucleotide sequence
XX
XX
SQ Sequence 158 BP; 41 A; 50 C; 21 G; 46 T; 0 U; 0 Other;
XX
XX
Query Match 80.7%; Score 24.2; DB 5; Length 158;
Best Local Similarity 89.7%; Pred. No. 1.5;
Matches 26; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
OY 1 CATGCTTGTCACCGTTCGCTTGTTC 29
DB 25 CATGCTTGTCACCGTTCGCTTGTTC 53
XX
XX
RESULT 8
ABO75351/c
ID ABO75351 standard; cDNA; 1304 BP.
XX
XX
AC ABO75351;
XX
XX
DT 05-NOV-2002 (first entry)
XX
XX
DE Human lung specific nucleic acid sequence SEQ ID NO:90.
XX
XX
KM Human; lung; lung specific nucleic acid; LSNA; lung specific protein;
KM LSP; cytostatic; gene therapy; vaccine; metastasis; lung cancer;
KM squamous cell carcinoma; gene; ss.
XX
XX
OS Homo sapiens.
XX
XX
PN WO200264788-A2.

```

```

XX
XX
PD 22-AUG-2002.
XX
XX
PF 20-NOV-2001; 2001WO-US045080.
XX
XX
PR 20-NOV-2000; 2000US-0252054P.
XX
XX
PA (DIAD-) DIADEXUS INC.
XX
XX
PI Macina RA, Recipon H, Chen S, Sun Y, Liu C;
XX
XX
DR WPI; 2002-657601/70.
XX
XX
PT New lung specific nucleic acid useful in gene therapy or as vaccines for
PT treating lung cancer (e.g. squamous cell carcinoma) or non-carcinous lung
PT diseases, as well as for diagnosing, monitoring or staging these
PT diseases.
XX
XX
PS Claim 1; Page 206; 282pp; English.
XX
XX
CC The present invention describes an isolated lung specific nucleic acid
CC (LSNA) comprising a sequence that: (a) encodes any of the 93 amino acid
CC sequences comprising 17 - 733 amino acids, given in ABP52873 to ABP52965;
CC (b) comprises any of 115 sequences comprising 148 - 3193 base pairs (bp);
CC given in ABO75262 to ABO75376; (c) selectively hybridizes to (a) or (b);
CC or (d) has 60 % sequence identity to (a) or (b). LSNA and lung specific
CC protein (LSP) sequences have cytostatic activity and can be used in gene
CC therapy and vaccines. LSNA and LSPs are useful for diagnosing and
CC monitoring the presence and metastases of lung cancer in a patient. An
CC antibody that specifically binds to an LSP can be used for determining
CC the presence of an LSP in a sample, as well as for treating a patient
CC with lung cancer, particularly by inducing an immune response against the
CC lung cancer cell expressing the LSNA or LSPs. In particular, these LSNA
CC and LSPs are useful for identifying, diagnosing, monitoring, staging,
CC imaging and treating lung cancer (e.g. squamous cell carcinoma) and non-
CC carcinous disease states in lung
XX
XX
SQ Sequence 1304 BP; 352 A; 347 C; 314 G; 291 T; 0 U; 0 Other;
XX
XX
Query Match 73.3%; Score 22; DB 6; Length 1304;
Best Local Similarity 83.3%; Pred. No. 16;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
XX
OY 1 CATGCTTGTCACCGTTCGCTTGTCCA 30
DB 1155 CTTTCTTGTCACCGTTCGCTTGTCCA 1126
XX
XX
RESULT 9
ABZ13819
ID ABZ13819 standard; DNA; 1542 BP.
XX
XX
AC ABZ13819;
XX
XX
DT 21-JAN-2003 (first entry)
XX
XX
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 1624.
XX
XX
KM Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX
XX
OS Arabidopsis thaliana.
XX
XX
PN WO200216655-A2.
XX
XX
PD 28-FEB-2002.
XX
XX
PF 24-AUG-2001; 2001WO-US026685.
XX
XX
PR 24-AUG-2000; 2000US-0227466P.
XX
XX
PR 26-JAN-2001; 2001US-0264647P.
XX
XX
PR 22-JUN-2001; 2001US-0300111P.
XX
XX
PA (SCRI ) SCRIPPS RES INST.

```

PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Harper JF, Kreps J, Wang X, Zhu T;
XX
DR WPI; 2002-304127/34.
XX
PT Identifying a stress condition to which a plant cell has been exposed and
XX producing plants with increased tolerance to these abiotic stresses.
PS
XX Claim 14; SEQ ID NO 1624; 577pp + Sequence Listing; English.
XX
CC The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising: (a) contacting nucleic acid
CC representative of expressed polynucleotides in the plant cell with an
CC array or probes representative of the plant cell genome; and (b)
CC detecting a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants
CC with increased tolerance to abiotic stress. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used
CC in methods of the invention. Note: The sequence data for this patent is
CC not represented in the printed specification but is based on sequence
CC information supplied to Derwent by the European Patent Office
XX
SQ Sequence 1542 BP; 345 A; 382 C; 370 G; 445 T; 0 U; 0 Other;
XX
Query Match 68.0%; Score 20.4; DB 6; Length 1542;
Best Local Similarity 80.0%; Pred. No. 81;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
XX
QY 1 CATGCTGTTCACCGTTCGTCCTTGCCA 30
DB 237 CATGATTGTTCACCAACATCGTCGTTCCA 266
XX
RESULT 10
ACA44013
ID ACA44013 standard; DNA; 2397 BP.
XX
AC ACA44013;
XX
DT 19-JUN-2003 (first entry)
XX
DE Prokaryotic essential gene #25670.
XX
KM Antisense; ds; prokaryotic essential gene; cell proliferation;
XX drug design; gene.
XX
OS Pseudomonas putida.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
XX
PR 06-SEP-2001; 2001US-00948993.
XX
PR 25-OCT-2001; 2001US-0342923P.
XX
PR 08-FEB-2002; 2002US-00072851.
XX
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haseelbeck R, Ohlsen KL, Zyskind JW;
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Foreyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
XX
DR P-PSDB; ABU40143.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.

XX
PS Claim 14; SEQ ID NO 31883; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2397 BP; 415 A; 825 C; 781 G; 376 T; 0 U; 0 Other;
XX
Query Match 67.3%; Score 20.2; DB 8; Length 2397;
Best Local Similarity 88.0%; Pred. No. 16+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
QY 1 CATGCTGTTCACCGTTCGTCCTTG 25
DB 516 CATGCTGTTCACCGTTCGCTTG 540
XX
RESULT 11
ADJ40785
ID ADJ40785 standard; cDNA; 2000 BP.
XX
AC ADJ40785;
XX
DT 06-MAY-2004 (first entry)
XX
DE Plant cDNA #1785.
XX
KM Plant; gene; ss; transcription; plant genome augmentation; cereal;
XX soybean; alfalfa; sunflower; canola; cotton; peanut; tobacco; sugar beet;
XX maize; barley; sorghum; rice; wheat; crop plant; insecticide resistance;
XX stress tolerance; salt tolerance; cold tolerance; drought tolerance;
XX plant nutrition; apical dominance; dwarfism; early flowering; antiviral;
XX antifungal.
XX
OS Eukaryota.
XX
PN US2004016025-A1.
XX
PD 22-JAN-2004.
XX
PF 26-SEP-2002; 2002US-00260238.
XX
PR 26-SEP-2001; 2001US-0325277P.
XX
PR 26-SEP-2001; 2001US-0325448P.

```

PR 04-APR-2002; 2002US-0370620P.
XX
XX (BUDW/) BUDWORTH P.
PA (MOUG/) MOUGHAMER T.
PA (BRIG/) BRIGGS S P.
PA (COOP/) COOPER B.
PA (GLAZ/) GLAZEBROOK J.
PA (GOFF/) GOFF S A.
PA (KATG/) KATAGIRI F.
PA (KREP/) KREPS J.
PA (PROV/) PROVART N.
PA (RICK/) RICE D.
PA (ZHUT/) ZHU T.
XX
XX Budworth P, Moughamer T, Briggs SP, Cooper B, Glazebrook J;
PI Goff SA, Katagiri F, Kreps J, Provart N, Ricke D, Zhu T;
XX
XX WPI; 2004-190374/18.
XX
XX New rice promoter, useful for manipulating crop plants to alter or
PT improve phenotypic characteristics, e.g. produce large quantities of oil
PT or proteins, resistance to insecticides, virus or fungi, stress tolerance
PT or high nutritional value.
XX
XX Claim 26; SEQ ID NO 1785; 230pp; English.
XX
XX The invention relates to plant nucleotide sequences that direct seed-,
CC leaf- and/or stem-, panicle-, root- or pollen-specific or -preferential
CC or constitutive transcription of an operatively linked nucleic acid
CC segment. The invention also relates to a method for augmenting a plant
CC genome and a method of identifying a gene, where its expression is
CC altered in the seed, leaf, stem, panicle, pollen, root or is constitutive
CC in a plant cell. The plant is a cereal, e.g. soybean, alfalfa, sunflower,
CC canola, cotton, peanut, tobacco or sugar beet, preferably maize, barley,
CC sorghum, rice or wheat. The polynucleotides and the polypeptides they
CC encode are useful for manipulating crop plants to alter or improve
CC phenotypic characteristics, to produce large quantities of oil or
CC proteins, to incur resistance to insecticides, viruses or fungi, and to
CC incur stress tolerance (e.g. salt, cold or drought) to ensure the plants
CC have a high nutritional value with reduced apical dominance or dwarfism,
CC early flowering or altered metabolic pathways. This sequence represents a
CC plant nucleic acid of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification but was obtained in
CC electronic format directly from USPTO at seqdata.uspto.gov/sequence.html.
XX
XX Sequence 2000 BP; 552 A; 438 C; 444 G; 566 T; 0 U; 0 Other;
SQ
XX
XX Query Match 66.7%; Score 20; DB 12; Length 2000;
XX Best Local Similarity 82.1%; Pred. No. 1.2e+02;
XX Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 3 TGCTTGTCACCGTCGTCGTTGTTCCA 30
DB 1179 TACATGTTTAAACCGTCGTCCTTATTCAA 1206

```

```

PD 03-JAN-2003.
XX
XX 22-JUN-2001; 2001WO-1B001105.
XX
XX 22-JUN-2001; 2001WO-1B001105.
XX
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Quan S, Rao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX
XX WPI; 2003-175290/17.
XX
XX Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
XX
XX Claim 27; SEQ ID NO 5386; 899pp; English.
XX
XX The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.
XX
XX Sequence 2000 BP; 605 A; 377 C; 326 G; 690 T; 0 U; 2 Other;
SQ
XX
XX Query Match 65.3%; Score 19.6; DB 8; Length 2000;
XX Best Local Similarity 84.6%; Pred. No. 1.8e+02;
XX Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 5 CTGTTCAACCGTCGTCGTTGTTCCA 30
DB 974 CATGTTTAAACCGTCGTCCTTATTCAA 999

```

```

RESULT 12
ADA72061
ID ADA72061 standard; DNA; 2000 BP.
XX
XX ADA72061;
AC
XX
XX 20-NOV-2003 (first entry)
DT
XX
XX Rice gene, SEQ ID 5386.
DE
XX
XX Plant; bacterial infection; fungal infection; viral infection; rice;
KM gene; ds.
XX
XX Oryza sativa.
OS
XX WO2003000898-A1.
XX
XX PN
XX
XX PT

```

```

RESULT 13
ADA71994/C
ID ADA71994 standard; DNA; 2000 BP.
XX
XX ADA71994;
AC
XX
XX 20-NOV-2003 (first entry)
DT
XX
XX Rice gene, SEQ ID 5319.
DE
XX
XX Plant; bacterial infection; fungal infection; viral infection; rice;
KM gene; ds.
XX
XX Oryza sativa.
OS
XX WO2003000898-A1.
XX
XX 03-JAN-2003.
PD
XX
XX 22-JUN-2001; 2001WO-1B001105.
PF
XX
XX 22-JUN-2001; 2001WO-1B001105.
PR
XX
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Quan S, Rao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX
XX WPI; 2003-175290/17.
XX
XX Identifying at least one gene involved in plant resistance or response to
PT

```


PT pathogenic infection for conferring resistance or tolerance to a plant to
 PT bacterial, fungal or viral infection by determining or detecting plant
 gene expression.
 XX
 PS Claim 27, SEQ ID NO 5319; 899bp; English.
 XX
 CC The present invention relates to a method (M1) for identifying genes
 CC involved in plant resistance or response to pathogenic infection. M1
 CC comprises identifying a gene whose expression is significantly altered in
 CC the incompatible interaction of plant gene expression relative to
 CC expression of the gene in an uninfected plant, in a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding incompatible or compatible interaction. (M1) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.
 XX
 SQ Sequence 2000 BP; 638 A; 400 C; 347 G; 615 T; 0 U; 0 Other;
 XX
 Query Match 65.3%; Score 19.6; DB 8; Length 2000;
 Best Local Similarity 84.6%; Pred. No. 1.8e+02;
 Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 OY 5 CTGTTCAACCGTTCGCTTGTTCGA 30
 DB 150 CATGTTAACCGTTCGCTTATTCA 125

RESULT 14
 ADA71736/c
 ID ADA71736 standard; DNA; 2000 BP.
 XX
 AC ADA71736;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Rice gene, SEQ ID 5061.
 XX
 KM Plant; bacterial infection; fungal infection; viral infection; rice;
 KM gene; ds.
 OS
 OS Oryza sativa.
 XX
 PN WO2003000898-A1.
 XX
 PD 03-JAN-2003.
 XX
 PF 22-JUN-2001; 2001WO-IB001105.
 XX
 PR 22-JUN-2001; 2001WO-IB001105.
 XX
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX
 PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
 PI Katagiri F, Quan S, Tao Y, Whitman S, Xie Z, Zhu T, Zou G;
 PI WPI; 2003-175290/17.
 DR
 PT Identifying at least one gene involved in plant resistance or response to
 PT pathogenic infection for conferring resistance or tolerance to a plant to
 PT bacterial, fungal or viral infection by determining or detecting plant
 gene expression.
 XX
 PS Claim 27, SEQ ID NO 5061; 899bp; English.
 XX
 CC The present invention relates to a method (M1) for identifying genes
 CC involved in plant resistance or response to pathogenic infection. M1
 CC comprises identifying a gene whose expression is significantly altered in
 CC the incompatible interaction of plant gene expression relative to
 CC expression of the gene in an uninfected plant, in a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding incompatible or compatible interaction. (M1) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC useful for conferring resistance to resistance or tolerance to a plant to

CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.
 XX
 SQ Sequence 2000 BP; 601 A; 422 C; 306 G; 671 T; 0 U; 0 Other;
 XX
 Query Match 65.3%; Score 19.6; DB 8; Length 2000;
 Best Local Similarity 84.6%; Pred. No. 1.8e+02;
 Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 OY 5 CTGTTCAACCGTTCGCTTGTTCGA 30
 DB 291 CATGTTAACCGTTCGCTTATTCA 266

RESULT 15
 ADA72530/c
 ID ADA72530 standard; DNA; 2000 BP.
 XX
 AC ADA72530;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Rice gene, SEQ ID 5855.
 XX
 KM Plant; bacterial infection; fungal infection; viral infection; rice;
 KM gene; ds.
 OS
 OS Oryza sativa.
 XX
 PN WO2003000898-A1.
 XX
 PD 03-JAN-2003.
 XX
 PF 22-JUN-2001; 2001WO-IB001105.
 XX
 PR 22-JUN-2001; 2001WO-IB001105.
 XX
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX
 PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
 PI Katagiri F, Quan S, Tao Y, Whitman S, Xie Z, Zhu T, Zou G;
 PI WPI; 2003-175290/17.
 DR
 PT Identifying at least one gene involved in plant resistance or response to
 PT pathogenic infection for conferring resistance or tolerance to a plant to
 PT bacterial, fungal or viral infection by determining or detecting plant
 gene expression.
 XX
 PS Claim 27, SEQ ID NO 5855; 899bp; English.
 XX
 CC The present invention relates to a method (M1) for identifying genes
 CC involved in plant resistance or response to pathogenic infection. M1
 CC comprises identifying a gene whose expression is significantly altered in
 CC the incompatible interaction of plant gene expression relative to
 CC expression of the gene in an uninfected plant, in a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding incompatible or compatible interaction. (M1) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.
 XX
 SQ Sequence 2000 BP; 549 A; 413 C; 443 G; 588 T; 0 U; 7 Other;
 XX
 Query Match 65.3%; Score 19.6; DB 8; Length 2000;
 Best Local Similarity 84.6%; Pred. No. 1.8e+02;
 Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 OY 5 CTGTTCAACCGTTCGCTTGTTCGA 30
 DB 27 CATGTTAACCGTTCGCTTATTCA 2

Sat Sep 17 09:06:39 2005

us-10-713-381-1_copy_1179_1208.rng

Page 8

Search completed: September 15, 2005, 21:33:14
Job time : 18.5242 secs

Sat Sep 17 09:06:40 2005

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 14, 2005, 19:14:19 : Search time 4.53786 Seconds

(without alignments)
10817.505 Million cell updates/sec

Title: US-10-713-381-1_COPY_1179_1208

Perfect score: 30
Sequence: 1 catgctgttcacacgctcgtcgttcctca 30

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/prodata/1/ina/5A COMB.seq:*
- 2: /cgn2_6/prodata/1/ina/5B COMB.seq:*
- 3: /cgn2_6/prodata/1/ina/6A COMB.seq:*
- 4: /cgn2_6/prodata/1/ina/6B COMB.seq:*
- 5: /cgn2_6/prodata/1/ina/PCUS COMB.seq:*
- 6: /cgn2_6/prodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	100.0	1394	3 US-08-880-499-1	Sequence 1, Appl
2	30	100.0	1394	3 US-08-880-499-2	Sequence 2, Appl
3	19.4	64.7	2243	1 US-07-937-609-15	Sequence 15, Appl
4	19.4	64.7	2243	3 US-08-029-170-15	Sequence 15, Appl
5	19.4	64.7	2243	4 US-09-443-745-15	Sequence 15, Appl
6	19	63.3	45571	4 US-09-949-016-16262	Sequence 16262, A
7	18.8	62.7	7812	3 US-09-368-590-1	Sequence 1, Appl
8	18.8	62.7	8756	4 US-09-949-016-1438	Sequence 1438, Ap
9	18.8	62.7	9541	3 US-08-961-527-132	Sequence 132, App
10	18.8	62.7	112874	4 US-09-949-016-13180	Sequence 13180, A
11	18.6	62.0	549	4 US-09-270-767-29480	Sequence 29480, A
12	18.6	62.0	1120	4 US-09-270-767-13492	Sequence 13492, A
13	18.4	61.3	1260	4 US-09-248-796A-3293	Sequence 3293, Ap
14	18.2	60.7	2710	4 US-09-573-080A-7	Sequence 7, Appl
15	18	60.0	601	4 US-09-949-016-20052	Sequence 20052, A
16	18	60.0	601	4 US-09-949-016-20053	Sequence 20053, A
17	18	60.0	889	4 US-08-956-171E-1150	Sequence 1190, Ap
18	18	60.0	889	4 US-08-781-986A-1190	Sequence 1190, Ap
19	18	60.0	1591	4 US-08-728-956-3	Sequence 3, Appl
20	18	60.0	111509	4 US-09-949-016-17379	Sequence 17379, A
21	18	60.0	254366	4 US-09-832-871-3	Sequence 3, Appl
22	17.8	59.3	393	4 US-09-583-110-1388	Sequence 1388, Ap
23	17.8	59.3	393	4 US-09-107-433-743	Sequence 743, App
24	17.8	59.3	499	4 US-09-270-767-1993	Sequence 1993, Ap
25	17.8	59.3	499	4 US-09-270-767-17275	Sequence 17275, A
26	17.8	59.3	601	4 US-09-949-016-142937	Sequence 142937, A
27	17.8	59.3	601	4 US-09-949-016-204029	Sequence 204029, A

C 28	17.8	59.3	601	4 US-09-949-016-204030	Sequence 204030, A
C 29	17.8	59.3	978	3 US-08-858-207A-174	Sequence 174, App
C 30	17.8	59.3	999	3 US-08-861-527-193	Sequence 193, App
C 31	17.8	59.3	2009	1 US-07-958-222A-1	Sequence 1, Appl
C 32	17.8	59.3	8073	4 US-09-949-016-12327	Sequence 12327, A
C 33	17.8	59.3	8074	4 US-09-949-016-16645	Sequence 16645, A
C 34	17.8	59.3	9472	1 US-08-325-547-9	Sequence 9, Appl
C 35	17.8	59.3	40505	4 US-09-949-016-13439	Sequence 13439, A
C 36	17.8	59.3	72843	4 US-09-949-016-12574	Sequence 12574, A
C 37	17.8	59.3	141115	4 US-09-949-016-17490	Sequence 17490, A
C 38	17.8	59.3	183202	4 US-09-949-016-13614	Sequence 13614, A
C 39	17.8	59.3	265038	4 US-09-949-016-15779	Sequence 15779, A
C 40	17.8	59.3	276237	4 US-09-949-016-17504	Sequence 17504, A
C 41	17.4	58.0	331	4 US-09-640-211A-169	Sequence 169, App
C 42	17.4	58.0	601	4 US-09-949-016-178197	Sequence 178197, A
C 43	17.4	58.0	719	4 US-09-107-433-820	Sequence 820, App
C 44	17.4	58.0	789	4 US-09-252-991A-11034	Sequence 11034, A
C 45	17.4	58.0	813	4 US-09-252-991A-5455	Sequence 5455, Ap

ALIGNMENTS

RESULT 1
US-08-880-499-1Sequence 1, Application US/08880499
Patent No. 6037523

GENERAL INFORMATION:

APPLICANT: Albersson, Marc C.

APPLICANT: Fox, Tim W.

APPLICANT: Carl, Gary A.

APPLICANT: Huffman, Gary A.

APPLICANT: Kendall, Timmy L.

TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION

TITLE OF INVENTION: AND METHOD OF USING SAME

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.

STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.

STREET: Box 1000

CITY: Johnston

STATE: Iowa

COUNTRY: USA

ZIP: 50131

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

FILING DATE: CONCURRENTLY HERewith

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Sweeney, Patricia A.

REGISTRATION NUMBER: 32,733

REFERENCE/DOCKET NUMBER: 0578

TELECOMMUNICATION INFORMATION:

TELEPHONE: (515) 248-4800

TELEFAX: (515) 248-4844

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1394 base pairs

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-880-499-1

Query Match 100.0%; Score 30; DB 3; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0.0009;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/029,170
; FILING DATE: 19930310
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/937,609
; FILING DATE: 02-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/928,033
; FILING DATE: 11-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/861,769
; FILING DATE: 01-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/831,248
; FILING DATE: 07-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 40399/166 NIHD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2243 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: rat brain CCKB receptor
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 136..1494
; US-08-029-170-15

Query Match      64.7%; Score 19.4; DB 3; Length 2243;
Best Local Similarity 79.3%; Pred. No. 49;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      1 CATGCTGTTCACCGCTGCTGTTGCC 29
Db      5 CCGCTTGCTCAACTGCTGTTGTTTC 33

RESULT 5
US-09-443-745-15
; Sequence 15, Application US/09443745
; Patent No. 6706493
; GENERAL INFORMATION:
; APPLICANT: MANK, Stephen A.
; TITLE OF INVENTION: CLONING AND FUNCTIONAL EXPRESSION OF
; TITLE OF INVENTION: CHOLECYSTOKININ RECEPTOR-ENCODING DNA
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; City: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/443,745
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
```

```

; APPLICATION NUMBER: 08/029,170
; FILING DATE: 10-MAR-1993
; APPLICATION NUMBER: US 07/937,609
; FILING DATE: 02-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/928,033
; FILING DATE: 11-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/861,769
; FILING DATE: 01-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/831,248
; FILING DATE: 07-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 40399/166 NIHD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2243 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: rat brain CCKB receptor
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 136..1494
; US-09-443-745-15
```

```

Query Match      64.7%; Score 19.4; DB 4; Length 2243;
Best Local Similarity 79.3%; Pred. No. 49;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      1 CATGCTGTTCACCGCTGCTGTTGCC 29
Db      5 CCGCTTGCTCAACTGCTGTTGTTTC 33
```

```

RESULT 6
US-09-949-016-16262
; Sequence 16262, Application US/09949016
; Patent No. 6812338
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16262
; LENGTH: 45571
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-16262

Query Match      63.3%; Score 19; DB 4; Length 45571;
Best Local Similarity 81.5%; Pred. No. 1,3e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

Qy	3	TGCTGTCAACCGTTCGTCCTGCC	29
Db	38119	TGTTTGTTCATTGCTTCCTTCCTTCC	38145

```

RESULT 7
US-09-368-590-1/C
; Sequence 1, Application US/09368590
; Patent No. 6187563
; GENERAL INFORMATION:
; APPLICANT: Solimena, Michele
; TITLE OF INVENTION: INTERACTING POLYPEPTIDES FOR
; FILE REFERENCE: 101918-200 (OCR-941)
; CURRENT APPLICATION NUMBER: US/09/368, 590
; CURRENT FILING DATE: 1999-08-04
; EARLIER APPLICATION NUMBER: 60/095, 657
; EARLIER FILING DATE: 1998-08-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 7812
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(6879)
; NAME/KEY: unsure
; LOCATION: (100)...(1102)
; NAME/KEY: unsure
; LOCATION: (1021)...(1023)
; NAME/KEY: unsure
; LOCATION: (2266)...(2268)
US-09-368-590-1

```

	Query Match	Similarity	Score	DB	Length
Best Local	841	76.7%	1.1e+02	812	7812
Matches	23	Conservative	0	Mismatches	7
				Indels	0
				Gaps	0

RESULT 8
 US-09-949-016-1438/C
 Sequence 1438, Application US/09949016
 Patent No. 6812339
 GENERAL INFORMATION:
 APPLICANT: VENTER, J. Craig et al.
 TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 FILE REFERENCE: CL001307
 CURRENT APPLICATION NUMBER: US/09/949,016
 CURRENT FILING DATE: 2000-04-14
 PRIOR APPLICATION NUMBER: 60/241,755
 PRIOR FILING DATE: 2000-10-20
 PRIOR APPLICATION NUMBER: 60/237,768
 PRIOR FILING DATE: 2000-10-03
 PRIOR APPLICATION NUMBER: 60/231,498
 PRIOR FILING DATE: 2000-09-08
 NUMBER OF SEQ ID NOS: 207012
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 1438
 LENGTH: 8756
 TYPE: DNA
 ORGANISM: Human
 US-09-949-016-1438

Query Match	62.7%	Score 18.8;	DB 4;	Length 8756;
Best Local Similarity	76.7%	Pred. No. 1.2e+02;		
Matches 23; Conservative	0;	Mismatches 7;	Indels 0;	Gaps 0;

QY 1 CATGCTGTCAACCGTTGCTGTTC 30
||| ||| ||| ||| ||| ||| |||
Db 1912 CATGCTGTGCACAGGTGCTGTCC 1883

```

1  RESULT 9
2  US-08-961-527-132
3  / Sequence 132, Application US/08961527
4  / Patent No. 6420135
5  / GENERAL INFORMATION:
6  / APPLICANT: Charles Kunsch
7  / TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
8  / NUMBER OF SEQUENCES: 391
9  / CORRESPONDENCE ADDRESS:
10 / ADDRESSEE: Human Genome Sciences, Inc.
11 / STREET: 9410 Key West Avenue
12 / CITY: Rockville
13 / STATE: Maryland
14 / COUNTRY: USA
15 / ZIP: 20850
16 /
17 / COMPUTER READABLE FORM:
18 / MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
19 / COMPUTER: HP Vectra 486/33
20 / OPERATING SYSTEM: MSDOS version 6.2
21 / SOFTWARE: ASCII Text
22 /
23 / CURRENT APPLICATION DATA:
24 / APPLICATION NUMBER: US/08/961,527
25 / FILING DATE:
26 / CLASSIFICATION: 424
27 / PRIOR APPLICATION DATA:
28 / APPLICATION NUMBER:
29 / FILING DATE:
30 /
31 / ATTORNEY/AGENT INFORMATION:
32 / NAME: Brookes, A. Anders
33 / REGISTRATION NUMBER: 36,733
34 / REFERENCE/DOCKET NUMBER: PB340P1
35 / TELECOMMUNICATION INFORMATION:
36 / TELEPHONE: (301) 309-8504
37 / TELEFAX: (301) 309-8512
38 / INFORMATION FOR SEQ ID NO: 132:
39 / SEQUENCE CHARACTERISTICS:
40 / LENGTH: 9541 base pairs
41 / TYPE: nucleic acid
42 / STRANDEDNESS: double
43 / TOPOLOGY: linear
44 /
45 / US-08-961-527-132

```

	Query Match	Similarity	Score 18.8	DB 3	Length 9541
Best Local Match	76.7%	Pred. No. 1.2e+02			
Matches 23	Conservative 0	Mismatches 7		Indels 0	Gaps 0
Qy	1	CATGCTGTTCACCGCTTCGCTTGTTC	30		
db	8049	CAGGCTGTTCACCAACGTTCCATTTTTC	8078		

RESULT 10
 US-09-949-016-13180/C
 ; Sequence 13180, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; TITLE OR INVENTION: WITH HUMAN DISEASE. METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949.016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ;

Sat Sep 17 09:06:40 2005

```

; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13180
; LENGTH: 112874
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(112874)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13180

Query Match          62.7%; Score 18.8; DB 4; Length 112874;
Best Local Similarity 76.7%; Pred. No. 1.9e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CATGCTTGTCAACCGTTCGTTGTTCCA 30
DB 41167 CATGCTTGTCAACAGGTCTGCTCCCA 41138

RESULT 11
US-09-270-767-29480/C
; Sequence 29480, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 29480
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-29480

Query Match          62.0%; Score 18.6; DB 4; Length 549;
Best Local Similarity 84.0%; Pred. No. 84;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 TGCTTGTCAACCGTTCGTTGTT 27
DB 223 TGCTTGTGAAGAGTCTCTTGT 199

RESULT 12
US-09-270-767-13492/C
; Sequence 13492, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 13492
; LENGTH: 1120
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-13492

Query Match          62.0%; Score 18.6; DB 4; Length 1120;
Best Local Similarity 84.0%; Pred. No. 96;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 TGCTTGTCAACCGTTCGTTGTT 27
DB 223 TGCTTGTGAAGAGTCTCTTGT 199

```

```

RESULT 13
US-09-248-796A-3293
; Sequence 3293, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 3293
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-3293

Query Match          61.3%; Score 18.4; DB 4; Length 1260;
Best Local Similarity 78.6%; Pred. No. 1.2e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 TGCTTGTCAACCGTTCGTTGTTCCA 30
DB 903 TGCTTGTCAAGTGTGTTCCGCCA 930

```

```

RESULT 14
US-09-573-080A-7/C
; Sequence 7, Application US/09573080A
; Patent No. 6828097
; GENERAL INFORMATION:
; APPLICANT: JOAN, KNOLL
; APPLICANT: ROGAN, PETER
; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATING
; FILE REFERENCE: 30307
; CURRENT APPLICATION NUMBER: US/09/573,080A
; CURRENT FILING DATE: 2000-05-16
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 7
; LENGTH: 2710
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: repeat region
; LOCATION: (1)..(2710)
; OTHER INFORMATION: charlie3
; NAME/KEY: misc feature
; OTHER INFORMATION: n is a, c, g or t
; PUBLICATION INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS: Jurka, J; Malchukiewicz, J; Milosavljevic, A
; TITLE: Prototypic sequences for human repetitive DNA
; JOURNAL: Journal of Molecular Evolution
; VOLUME: 35
; ISSUE: 4
; PAGES: 286-291
; DATE: 1992-10-
; DATABASE ACCESSION NUMBER: Database for repetitive elements (repbase)
; DATABASE ENTRY DATE: 1996-01-26
US-09-573-080A-7

Query Match          60.7%; Score 18.2; DB 4; Length 2710;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 20; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```

Qy 3 TGCTGTTCAACCGTTCGTTGTT 27
Db 1209 TGCTGTTTCACGAGTTCGTTKGTT 1185

RESULT 15
US-09-949-016-200252/c
; Sequence 200252, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 200252
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-200252

Query Match 60.0%; Score 18; DB 4; Length 601;
Best local similarity 80.8%; Pred. No. 1.6e+02;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 5 CTGTGTTCAACCGTTCGTTGTTGTTCA 30
Db 239 CTGTGTTAAACAGTTCCTGTTGTTCA 214

Search completed: September 15, 2005, 08:25:07
Job time : 7.53786 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: September 15, 2005, 07:10:11 ; Search time 19.5379 Seconds
(without alignment)
10230.248 Million cell updates/sec

Title: US-10-713-381-1_COPY_1179_1208

Perfect score: 30
Sequence: 1 catgctgttcaacgcttgcgtctgtcca 30

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7389322 seqs, 333185599 residues

Total number of hits satisfying chosen parameters: 14778644

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_NEW_PUB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	100.0	30	US-10-713-381-4	Sequence 4, Appli
2	30	100.0	255	US-10-713-381-9	Sequence 9, Appli
3	30	100.0	1394	US-10-713-381-1	Sequence 1, Appli
4	30	100.0	1394	US-10-713-381-2	Sequence 2, Appli
5	24.2	80.7	158	US-10-713-381-3	Sequence 3, Appli
6	22	73.3	1304	US-10-001-857-90	Sequence 90, Appli
7	21	70.0	1980090	US-10-719-993-6815	Sequence 6815, Ap

C	8	21	70.0	1980900	21	US-10-741-600-17676	Sequence 17676, A
9	20.4	68.0	1542	9	US-09-938-842A-1624	Sequence 1624, Ap	
10	20.4	68.0	1542	11	US-09-938-842A-1624	Sequence 1624, Ap	
11	20.2	67.3	2397	17	US-10-282-122A-31883	Sequence 31883, A	
12	20	66.7	2000	17	US-10-260-238-1785	Sequence 1785, Ap	
C	13	19.6	65.3	640	US-10-437-963-96639	Sequence 96639, A	
C	14	19.6	65.3	901	US-10-425-115-3534	Sequence 3534, A	
C	15	19.6	65.3	2000	US-10-260-238-1977	Sequence 1977, Ap	
C	16	19.4	64.7	2243	US-09-443-745-15	Sequence 15, Appli	
17	19.4	64.7	2256646	19	US-10-470-565-1	Sequence 1, Appli	
C	18	19	63.3	1041	US-10-425-115-100318	Sequence 100318, A	
C	19	63.3	1821	17	US-10-425-563-5	Sequence 5, Appli	
C	20	19	63.3	1821	US-10-972-789A-5	Sequence 5, Appli	
C	21	19	63.3	5671	10	US-09-764-872-729	Sequence 729, Ap
C	22	19	63.3	198522	13	US-10-087-192-244	Sequence 244, Ap
C	23	18.8	62.7	883	10	US-09-934-455-279	Sequence 279, Ap
C	24	18.8	62.7	883	17	US-10-225-068-183	Sequence 183, Ap
C	25	18.8	62.7	883	17	US-10-325-068A-971	Sequence 971, Ap
C	26	18.8	62.7	883	17	US-10-374-780A-381	Sequence 381, Ap
C	27	18.8	62.7	883	21	US-10-225-068-183	Sequence 183, Ap
C	28	18.8	62.7	883	22	US-10-225-068A-971	Sequence 971, Ap
C	29	18.8	62.7	1083	9	US-09-815-242-7281	Sequence 7281, Ap
C	30	18.8	62.7	1083	17	US-10-282-122A-22639	Sequence 22639, A
C	31	18.8	62.7	1083	18	US-10-335-977-2194	Sequence 2194, A
C	32	18.8	62.7	1125	18	US-10-335-977-2195	Sequence 2195, Ap
C	33	18.8	62.7	1856	19	US-10-437-963-40048	Sequence 40048, A
C	34	18.8	62.7	1859	21	US-10-472-928-2217	Sequence 2217, Ap
C	35	18.8	62.7	1902	9	US-09-815-242-9220	Sequence 9220, Ap
C	36	18.8	62.7	1902	17	US-10-282-122A-37788	Sequence 37788, A
C	37	18.8	62.7	2000	17	US-10-260-238-2535	Sequence 2535, Ap
C	38	18.8	62.7	2247	17	US-10-369-493-33483	Sequence 33483, A
C	39	18.8	62.7	2920	17	US-10-108-260A-1041	Sequence 1041, Ap
C	40	18.8	62.7	8788	22	US-10-756-149-64	Sequence 64, Appli
C	41	18.8	62.7	9541	8	US-08-961-527-132	Sequence 132, Ap
C	42	18.8	62.7	9541	17	US-10-158-844-132	Sequence 132, Ap
C	43	18.8	62.7	2165598	21	US-10-472-928-4979	Sequence 4979, Ap
C	44	18.6	62.0	375	9	US-09-974-300-3242	Sequence 3242, Ap
C	45	18.4	61.3	174	10	US-09-814-353-5230	Sequence 5230, Ap

ALIGNMENTS

RESULT 1
US-10-713-381-4
; Sequence 4, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 4
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-4

Query Match 100.0%; Score 30; DB 20; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.0065;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CATGCTGTTCACGCTTGCCTCTGTTCACA 30

Db 1 CATGCTGTTCACCGTTCGTCTGTTC 30

RESULT 2
US-10-713-381-9
; Sequence 9, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 255
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-9

Query Match 100.0%; Score 30; DB 20; Length 255;
Best Local Similarity 100.0%; Pred. No. 0.0086;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGCTGTTCACCGTTCGTCTGTTC 30
DB 39 CATGCTGTTCACCGTTCGTCTGTTC 68

RESULT 3
US-10-713-381-1
; Sequence 1, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-1

Query Match 100.0%; Score 30; DB 20; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGCTGTTCACCGTTCGTCTGTTC 30
DB 1179 CATGCTGTTCACCGTTCGTCTGTTC 1208

RESULT 4

US-10-713-381-2
; Sequence 2, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-2

Query Match 100.0%; Score 30; DB 20; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGCTGTTCACCGTTCGTCTGTTC 30
DB 1179 CATGCTGTTCACCGTTCGTCTGTTC 1208

RESULT 5
US-10-713-381-3
; Sequence 3, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 158
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-3

Query Match 80.7%; Score 24.2; DB 20; Length 158;
Best Local Similarity 89.7%; Pred. No. 2.2;
Matches 26; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CATGCTGTTCACCGTTCGTCTGTTC 29
DB 25 CATGCTGTTCACCGTTCGTCTGTTC 53

RESULT 6
US-10-001-857-90/c
; Sequence 90, Application US/10001857
; Publication No. US20020183500A1
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto

```
; APPLICANT: Recipon, Hervé
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; FILE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Pro
; FILE REFERENCE: DEX-0273
; CURRENT APPLICATION NUMBER: US/10/001,857
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/252,054
; PRIOR FILING DATE: 2000-11-20
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 90
; LENGTH: 1304
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-001-857-90

Query Match          73.3%; Score 22; DB 13; Length 1304;
Best Local Similarity 83.3%; Pred. No. 25;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CATGCTGTTCACCGTTCGCTTGTTCCA 30
Db 1155 CTTCTTGTTCCTCCCGCTCTCTTGTTCCA 1126

RESULT 7
US-10-719-993-6815/c
; Sequence 6815, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001496
; CURRENT APPLICATION NUMBER: US/10/719,993
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 53342
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6815
; LENGTH: 1980090
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1980090)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-719-993-6815

Query Match          70.0%; Score 21; DB 20; Length 1980090;
Best Local Similarity 82.8%; Pred. No. 1.7e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 ATGCTGTTCACCGTTCGCTTGTTCCA 30
Db 1948918 ATTCTTGCTGACAGTTCGCTTGTTCACCA 1948890

RESULT 8
US-10-741-600-17676/c
; Sequence 17676, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17676
```

```
; LENGTH: 1980090
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1980090)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-741-600-17676

Query Match          70.0%; Score 21; DB 21; Length 1980090;
Best Local Similarity 82.8%; Pred. No. 1.7e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 ATGCTGTTCACCGTTCGCTTGTTCCA 30
Db 1948918 ATTCTTGCTGACAGTTCGCTTGTTCACCA 1948890

RESULT 9
US-09-938-842A-1624
; Sequence 1624, Application US/0938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Joel
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1624
; LENGTH: 1542
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1624

Query Match          68.0%; Score 20.4; DB 9; Length 1542;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CATGCTGTTCACCGTTCGCTTGTTCCA 30
Db 237 CATGATTTTCACCAATGTCGCTTTCACCA 266

RESULT 10
US-09-938-842A-1624
; Sequence 1624, Application US/0938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Joel
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
```

;; PRIOR FILING DATE: 2001-06-22
;; NUMBER OF SEQ ID NOS: 5379
;; SEQ ID NO 1624
;; LENGTH: 1542
;; TYPE: DNA
;; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1624

Query Match 68.0%; Score 20.4; DB 11; Length 1542;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CATGCTGTTCACCGTTCGTTGTTCCA 30
DB 237 CATGATTGTTCACCAATCGTGGTTTCCA 266

RESULT 11
US-10-282-122A-31883
; Sequence 31883, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITTA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 7614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31883
; LENGTH: 2397
; TYPE: DNA
; ORGANISM: Pseudomonas putida
US-10-282-122A-31883

Query Match 67.3%; Score 20.2; DB 17; Length 2397;
Best Local Similarity 88.0%; Pred. No. 1.6e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CATGCTGTTCACCGTTCGTTGTTG 25
||||| ||||||| |||||

DB 516 CATGCTGTTCACCGTTCGCTTG 540

RESULT 12
US-10-260-238-1785
; Sequence 1785, Application US/10260238
; Publication No. US20040016025A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, Paul R.
; APPLICANT: Moughamer, Todd G.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Katagiri, Fumiyaki
; APPLICANT: Kreps, Joel
; APPLICANT: Provart, Nicholas
; APPLICANT: Rieke, Darrell
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION
; FILE REFERENCE: 60111-NP
; CURRENT APPLICATION NUMBER: US/10/260,238
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/325,448
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 6077
; SEQ ID NO 1785
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Oryza sativa
US-10-260-238-1785

Query Match 66.7%; Score 20; DB 17; Length 2000;
Best Local Similarity 82.1%; Pred. No. 1.9e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 TCGTGTTCACCGTTCGTTGTTCCA 30
DB 1179 TACATGTTTAAACGTTGCTTATTCAA 1206

RESULT 13
US-10-437-963-96639/C
; Sequence 96639, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 96639
; LENGTH: 640
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_94716C.1
US-10-437-963-96639

Query Match 65.3%; Score 19.6; DB 19; Length 640;

Best Local Similarity 84.6%; Pred. No. 2.4e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 5 CTGCTTCAACCGTTCGTTGTTCCA 30
Db 199 CATGTTTAAACCGTTCGTTATTCA 174

RESULT 14

US-10-425-115-3534/C
Sequence 3534, Application US/10425115
Publication No. US20040214272A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yinhua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 3534
LENGTH: 901
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_103225C.1
US-10-425-115-3534

Query Match 65.3%; Score 19.6; DB 20; Length 901;
Best Local Similarity 84.6%; Pred. No. 2.5e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CATGCTTGTCAACCGTTCGTTGT 26
Db 33 CACGCTTGTCAACCGTTCCTGTGT 8

RESULT 15

US-10-260-238-1977/C
Sequence 1977, Application US/10260238
Publication No. US20040016025A1
GENERAL INFORMATION:
APPLICANT: Budworth, Paul R.
APPLICANT: Moughamer, Todd G.
APPLICANT: Briggs, Steven P.
APPLICANT: Cooper, Bret
APPLICANT: Glazebrook, Jane
APPLICANT: Golf, Stephen A.
APPLICANT: Katsigiri, Fumiyaki
APPLICANT: Krepas, Joel
APPLICANT: Provart, Nicholas
APPLICANT: Rieke, Darrell
APPLICANT: Zhu, Tong
TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION
FILE REFERENCE: 6011-NP
CURRENT APPLICATION NUMBER: US/10/260,238
CURRENT FILING DATE: 2002-09-26
PRIOR APPLICATION NUMBER: US 60/325,448
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: US 60/325,277
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: US 60/370,620
PRIOR FILING DATE: 2002-04-04
NUMBER OF SEQ ID NOS: 6077
SEQ ID NO 1977
LENGTH: 2000
TYPE: DNA
ORGANISM: Oryza sativa
US-10-260-238-1977

Query Match 65.3%; Score 19.6; DB 17; Length 2000;
Best Local Similarity 84.6%; Pred. No. 2.7e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 5 CTGCTTCAACCGTTCGTTGTTCCA 30
Db 291 CATGTTTAAACCGTTCGTTATTCA 266

Search completed: September 15, 2005, 20:45:45
Job time: 26.5379 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 08:25:31 ; Search time 134.222 Seconds
(without alignments)
14440.280 Million cell updates/sec

Title: US-10-713-381-1_COPY_1239_1278

Perfect score: 40
Sequence: 1 aggatattcttctcccaaacatcttaccatgcaac 40

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenBank:.*
1: gb_ba.*
2: gb_ptg.*
3: gb_in.*
4: gb_cm.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_dl.*
9: gb_dr.*
10: gb_ro.*
11: gb_stb.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	40	6	AX224399 Sequence
2	40	100.0	50	6	AX224398 Sequence
3	40	100.0	158	6	AX224396 Sequence
4	40	100.0	255	6	AX224402 Sequence
5	40	100.0	1394	6	AX224394 Sequence
6	40	100.0	1394	6	AX224395 Sequence
7	40	100.0	1394	6	BD062176 Male t188
8	40	100.0	1394	6	BD062177 Male t188
9	40	100.0	1394	6	BD062177 Male t188
10	40	100.0	1394	6	BD062177 Male t188
11	40	100.0	1394	6	BD062177 Male t188
12	40	100.0	1394	6	BD062177 Male t188
13	40	100.0	1394	6	BD062177 Male t188
14	40	100.0	1394	6	BD062177 Male t188
15	40	100.0	1394	6	BD062177 Male t188
16	40	100.0	1394	6	BD062177 Male t188
17	40	100.0	1394	6	BD062177 Male t188
18	40	100.0	1394	6	BD062177 Male t188
19	40	100.0	1394	6	BD062177 Male t188

20	23	57.5	218224	10	AC134404	AC134404 Mus muscu
21	23	57.5	234163	2	AC115318	AC115318 Rattus no
22	23	57.5	242109	2	AC147523	AC147523 Oryzomys
23	23	57.5	319056	2	AC016640	AC016640 Homo sapi
24	22.8	57.0	164944	3	AC008195	AC008195 Drosophila
25	22.8	57.0	175781	3	AC008091	AC008091 Drosophila
26	22.8	57.0	235928	3	AE003736	AE003736 Drosophila
27	22.6	56.5	958968	2	AL156384	AL156384 Human DNA
28	22.6	56.5	130893	2	AL149638	AL149638 Human DNA
29	22.6	56.5	160246	2	AC009221	AC009221 Homo sapi
30	22.6	56.5	161817	10	AC132119	AC132119 Mus muscu
31	22.6	56.5	170973	9	AP000923	AP000923 Homo sapi
32	22.6	56.5	175110	9	AP002428	AP002428 Homo sapi
33	22.6	56.5	186457	2	AC129669	AC129669 Rattus no
34	22.6	56.5	196990	2	AC134013	AC134013 Rattus no
35	22.6	56.5	235953	2	AC130117	AC130117 Rattus no
36	22.4	56.0	112361	9	H839817	AL023656 Human DNA
37	22.2	55.5	21700	9	BX322645	BX322645 Human DNA
38	22.2	55.5	71561	2	AL928982_6	Continuation (7 of
39	22.2	55.5	73885	2	AC025307	AC025307 Homo sapi
40	22.2	55.5	111122	8	ATFL1C1	AL132976 Arabidops
41	22.2	55.5	150489	2	AC069491	AC069491 Homo sapi
42	22.2	55.5	152797	2	AC016856	AC016856 Homo sapi
43	22.2	55.5	153814	2	AP003180	AP003180 Homo sapi
44	22.2	55.5	157574	9	AL592293	AL592293 Human DNA
45	22.2	55.5	176084	2	AC113354	AC113354 Homo sapi

ALIGNMENTS

RESULT 1	AX224399	Sequence 6 from Patent WO0160997.	40 bp	DNA	linear	PAT 10-SEP-2001
LOCUS	AX224399					
DEFINITION	AX224399					
ACCESSION	AX224399					
VERSION	AX224399.1	GI:15554641				
KEYWORDS						
SOURCE						
ORGANISM						
Zeae maye						
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCB						
clade; Panicoidae; Andropogoneae; Zea.						
REFERENCE						
AUTHORS	Albertsen, M.C., Fox, T.W., Garnaat, C.W., Huffman, G. and Kendall, T.L.					
TITLE	Male tissue-preferred regulatory region and method of using same					
JOURNAL	Patent: WO 0160997-A 6 23-AUG-2001;					
PIONEER	HI-BRED INTERNATIONAL, INC. (US)					
FEATURES						
source						
1..40						
/organism="Zea maye"						
/mol_type="unassigned DNA"						
/db_xref="taxon:4577"						
ORIGIN						
Query Match	100.0%	Score 40;	DB 6;	Length 40;		
Best Local Similarity	100.0%;	Pred. No. 4.2e-05;				
Matches	40;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	1	AGGATACCTACTCTCCCAACATCCATCTTACTCATGCAAC	40			
Db	1	AGGATACCTACTCTCCCAACATCCATCTTACTCATGCAAC	40			
RESULT 2	AX224398	Sequence 5 from Patent WO0160997.	50 bp	DNA	linear	PAT 10-SEP-2001
LOCUS	AX224398					
DEFINITION	AX224398					
ACCESSION	AX224398					
VERSION	AX224398.1	GI:15554640				
KEYWORDS						
SOURCE						
ORGANISM						
Zeae maye						

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
1
Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
Patent: WO 0160997-A 5 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)

FEATURES
source
location/Qualifiers
1..50
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN

Query Match 100.0%; Score 40; DB 6; Length 50;
Best Local Similarity 100.0%; Pred. No. 4.1e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 40
|||||
11 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 50

Db

RESULT 3
AX224396 158 bp DNA linear PAT 10-SEP-2001
LOCUS
DEFINITION
Sequence 3 from Patent WO0160997.
AX224396
VERSION
AX224396.1 GI:15554638
KEYWORDS
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
1
Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
Patent: WO 0160997-A 3 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)

FEATURES
source
location/Qualifiers
1..158
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN

Query Match 100.0%; Score 40; DB 6; Length 158;
Best Local Similarity 100.0%; Pred. No. 3.6e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 40
|||||
86 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 125

Db

RESULT 4
AX224402 255 bp DNA linear PAT 10-SEP-2001
LOCUS
DEFINITION
Sequence 9 from Patent WO0160997.
AX224402
VERSION
AX224402.1 GI:15554644
KEYWORDS
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
1
Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
Patent: WO 0160997-A 9 23-AUG-2001;

AUTHORS
TITLE
JOURNAL

PIONEER HI-BRED INTERNATIONAL, INC. (US)

FEATURES
source
location/Qualifiers
1..255
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN

Query Match 100.0%; Score 40; DB 6; Length 255;
Best Local Similarity 100.0%; Pred. No. 3.4e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 40
|||||
99 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 138

Db

RESULT 5
AX224394 1394 bp DNA linear PAT 10-SEP-2001
LOCUS
DEFINITION
Sequence 1 from Patent WO0160997.
AX224394
VERSION
AX224394.1 GI:15554636
KEYWORDS
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
1
Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
Patent: WO 0160997-A 1 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)

FEATURES
source
location/Qualifiers
1..1394
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN

Query Match 100.0%; Score 40; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 2.8e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 40
|||||
1239 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 1278

Db

RESULT 6
AX224395 1394 bp DNA linear PAT 10-SEP-2001
LOCUS
DEFINITION
Sequence 2 from Patent WO0160997.
AX224395
VERSION
AX224395.1 GI:15554637
KEYWORDS
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
1
Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
Patent: WO 0160997-A 2 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)

FEATURES
source
location/Qualifiers
1..1394
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN

Query Match 100.0%; Score 40; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 2.8e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAACAATCCATCTTACTCATGCAAC 40
|||||
DB 1239 AGGATACCTACTCCCAACAATCCATCTTACTCATGCAAC 1278

RESULT 7
BD062176 1394 bp DNA linear PAT 27-AUG-2002
LOCUS Male tissue-preferred regulatory region and method of using same.
DEFINITION
ACCESSION BD062176 GI:22607781
VERSION JP 2001520523-A/1.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1394)
Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
Patent: JP 2001520523-A 2 30-OCT-2001;
PIONEER HI BRED INTERNATIONAL INC
COMMENT

REFERENCE
AUTHORS Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G.A. and Kendall,T.L.
TITLE Male tissue-preferred regulatory region and method of using same
JOURNAL PIONEER HI BRED INTERNATIONAL INC
PM JP 2001520523-A/1
PN JP 2001520523-A/1
PD 30-OCT-2001
PF 19-JUN-1998 JP 199504910
PR 23-JUN-1997 US 06/880499
PI MARC C ALBERTSEN,TIMOTHY W FOX,CARL W GARNAAAT,GARY A HUFFMAN,
PI TIMMY L KENDALL
PC C12N15/82,C12N15/29,C12N9/24,C12N9/22,C12N9/10,C12N9/00 PC
,C07K14/34,C12Q1/68,
PC A01H5/00
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers.
FEATURES
source 1..1394
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 40; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 2.8e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAACAATCCATCTTACTCATGCAAC 40
|||||
DB 1239 AGGATACCTACTCCCAACAATCCATCTTACTCATGCAAC 1278

RESULT 8
BD062177 1394 bp DNA linear PAT 27-AUG-2002
LOCUS Male tissue-preferred regulatory region and method of using same.
DEFINITION
ACCESSION BD062177 GI:22607782
VERSION JP 2001520523-A/2.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1394)
Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
Patent: JP 2001520523-A 2 30-OCT-2001;
PIONEER HI BRED INTERNATIONAL INC
COMMENT

REFERENCE
AUTHORS Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G.A. and Kendall,T.L.
TITLE Male tissue-preferred regulatory region and method of using same
JOURNAL PIONEER HI BRED INTERNATIONAL INC
PM JP 2001520523-A/2
PN JP 2001520523-A/2

PD 30-OCT-2001
PR 19-JUN-1998 JP 199504910
PR 23-JUN-1997 US 08/880499
PI MARC C ALBERTSEN,TIMOTHY W FOX,CARL W GARNAAAT,GARY A HUFFMAN,
PI TIMMY L KENDALL
PC C12N15/82,C12N15/29,C12N9/24,C12N9/22,C12N9/10,C12N9/00 PC
,C07K14/34,C12Q1/68,
PC A01H5/00
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers.
FEATURES
source 1..1394
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 40; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 2.8e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAACAATCCATCTTACTCATGCAAC 40
|||||
DB 1239 AGGATACCTACTCCCAACAATCCATCTTACTCATGCAAC 1278

RESULT 9
AF360356 3343 bp DNA linear PIN 12-MAY-2001
LOCUS Zea mays male fertility protein (Ms45) gene, complete cds.
DEFINITION
ACCESSION AF360356
VERSION AF360356.1 GI:14028756
KEYWORDS Zea mays
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 3343)
Fox,T.W., Trimmell,M.R. and Albertsen,M.C.
Cloning of Ms45, a gene required for male fertility from Zea mays
Unpublished
2 (bases 1 to 3343)
Fox,T.W., Trimmell,M.R. and Albertsen,M.C.
Direct Submision
Submitted (13-MAR-2001) Trait and Technology Development, Pioneer
Hi-Bred Intl. Inc., 7300 N.W. 62nd Ave., P.O. Box 1004, Johnston,
IA 50131-1004, USA
FEATURES
source 1..3343
/organism="Zea mays"
/mol_type="genomic DNA"
/culivar="B73"
/db_xref="taxon:4577"
/chromosome="9L"
<1392..>2942
/gene="Ms45"
/join(<1392..1768,1898..2182,2280..2447,2534..>2942)
/gene="Ms45"
/product="male fertility protein"
/join(1392..1768,1898..2182,2280..2447,2534..2942)
/gene="Ms45"
/codon_start=1
/product="male fertility protein"
/protein_id="AAK52489.1"
/db_xref="GI:14028757"
/translation="MSKNTLQWRBGRDGIQVYPHLPFAALMALALVADPPGSLAEV
DTRPVHILAPYGEVWGSWPRDNASRLRGRLEFGEVFGPSIFFDLQGRPYGLA
DGRVVRWGEAGMETFAVMNDWSEVCAAGVNSTRKQHKERKFCRPIGLRFGH
TGELVADVAYGLAVVQSGGVAASVAREADGDPTRFANDLDVHNSGVFFPDTSMRY
SKDHLNILLLEGEGTGRLLRYDPTSGVAVVLKGLVPRNGVQISDHQFLPSETTNC
RIMRYWLSGPRAGEVVFANLPGFPDYNRSNRQGFVAIDCRRPADEVFAKRWLR

TLTKFPLSLKVLTMKARMTLALLDDEGRVVELEDRGHEVMKLVSEVRGKR
LWIGTAAHHNIATIPYPLED"

ORIGIN

Query Match 100.0%; Score 40; DB 8; Length 3343;
Best Local Similarity 100.0%; Pred. No. 2,6e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAACAATCCATCTTACTCATCGAAC 40
|||||
Db 1239 AGGATACCTACTCCCAACAATCCATCTTACTCATCGAAC 1278

RESULT 10
AC135206/c 137327 bp DNA linear PLN 16-APR-2003
LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) chromosome 3 clone
OJ1041F02, complete sequence.

AC135206
AC135206.3 GI:27596977
HTG.

SOURCE

Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.

1 (bases 1 to 137327)
Wing, R.A., Yu, Y., Soderlund, C., Kim, H.-R., Rambo, T., Currie, J. and
Collura, K.

REFERENCE

Rice Genomic Sequence
Unpublished
2 (bases 1 to 137327)
Wing, R.A., Yu, Y., Soderlund, C., Kim, H.-R., Rambo, T., Sasaki, C.,
Currie, J., Collura, K. and Thompson, S.

REFERENCE

Direct Submission
Submitted (09-OCT-2002) Arizona Genomics Institute, University of
Arizona, 303 Forbes, Tucson, AZ 85721, USA

REFERENCE

Wing, R.A., Yu, Y., Soderlund, C., Kim, H.-R., Rambo, T., Currie, J.,
Currie, J., Collura, K. and Thompson, S.

REFERENCE

Direct Submission
Submitted (06-NOV-2002) Arizona Genomics Institute, University of
Arizona, 303 Forbes, Tucson, AZ 85721, USA

REFERENCE

Wing, R.A., Yu, Y., Soderlund, C., Kim, H.-R., Rambo, T., Currie, J.,
Collura, K. and Thompson, S.

REFERENCE

Direct Submission
Submitted (11-JAN-2003) Arizona Genomics Institute, University of
Arizona, 303 Forbes, Tucson, AZ 85721, USA

REFERENCE

Wing, R.A., Yu, Y., Soderlund, C., Chen, M., Kim, H., Rambo, T.,
Sasaki, C., Henry, D., Thompson, S., Simmons, J., Thurmond, S.K. and
Sun, S.

REFERENCE

Direct Submission
Submitted (16-APR-2003) Clemson University Genomics Institute,
Clemson University, 100 Jordan Hall, Clemson, SC 29634, USA

REFERENCE

On Jan 11, 2003 this sequence version replaced gi:24635891.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality
>30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by more than
one plasmid subclone; and the assembly was confirmed by
restriction digest. There are transposons located at area
29427-32874 and at area 19945-33212. Region 70511-70536 is covered
by a single subclone. Region 18678-18873 is a double stranded
single clone. Areas 69957-70761, 89836-90416, 98951-99009,
105316-105607, and 133052-133152 are covered by Monsanto only. The
nucleotide sequence of this BAC clone was generated by combining
Syngenta, Monsanto and Arizona Genomics Institute sequencing data.

FEATURES

1.137327
/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="genomic DNA"
/db_xref="taxon:39947"
/chromosome="3"
/clone="OJ1041F02"
/note="(japonica cultivar-group)"
complement(638..9495)
/gene="OJ1041F02.1"

CDS
complement(join(638..838,1141..1276,3132..3244,3326..3421,
3513..3727,3812..3934,4024..4054,4157..4216,5088..5243,
5322..5411,5484..5659,5808..5942,6035..6538,6622..6837,
7548..7593,7679..7815,8196..8328,9221..9495))
/gene="OJ1041F02.1"

/codon_start=1
/product="unknown protein"
/protein_id="AAP06844.1"
/db_xref="GI:29893590"

/translation="MAQSLLELLIQFLMPNDARRQAEQIRRLARDPQVPALVHLL
RTAKPNVROLAAVLLRKRTITSHWPKLPKAKASLKQALIDSTIDHSHLVRAQANV
VSIKAVVAPGAEPELPIFOGSGPOEDHREVALIFGSLTGTITGTPQSHNDL
OPIILKICLQDASSRVRITALKAVGSPFTEYVNDGVDVVKIPRDPVILANVROCLAN
GEEDVASIAFEIPDELIESPAFLGDSIRSYVQESLRYVCSNOLEINRQDIIQIISM
LVEKASFLKHKVLPILQWCPILTETABDGSIDLADNSAEVITDAINLPRH
VPPVLAESVSPRIINPKRYBAATVSLGVVSEGCERLKKLEDCLEKLVVLEALDOE
QMVKGASFAVGQFAHEHQPELISHYESVLCILNALDEPSEVSEKSYVALAACED
MGENILPYLDEPLMCRVLSQSPRNLOETGMSAIGVAAAQAFMFVYAEKLEEMK
GEMVLTNDEDI CARARATEVGIYMAVGRARMETILPPIEATISGVLDVSELRXY
THGFSNVAEILDSFAYLPRVAVLAASSCNDLGDGSAVIDDADSYNNGSGVSDSD
DVNDSPVRNLSMLTITADILTAITPAHADVLEKODILDTWNIYIKTRREDID
KEVNAQACTSLADIVRDCGFAIIEBYITRLADATILLRQSCCQGVSDGDDGID
HDEVLMADSDLLPAFAKVMGSEYDPITFKLFDLSMKPKASPHPDQTMVATLAEV
AOGMGAPISAVYDKIMPLVLEKLSSEATNRNAPCVGEMCKNGGAALAKYGDILH
GLHRFADSEPDPAARDNAAGILARMIWVQOSIPLNQIILPVPVITVIAFQVAVVSPN
ESDEVKTVAKAVASHLISVYGQMPILSLAPPAANALASPARNR"

12372..15503
/gene="OJ1041F02.2"
/note="unknown protein"
join(12372..12757,12858..13025,13130..13295,13746..13826,
13981..14130,14224..14354,14459..14558,15197..15503)
/gene="OJ1041F02.2"

/codon_start=1
/product="unknown protein"
/protein_id="AAP06845.1"
/db_xref="GI:29893591"

/translation="MPQGGGCRHGMVGEBAVGGGEBPVVPLRKRLSVDGAS
TAPRTICEGGEAGDITCDIYAAPLRSCSAKAMPAPPLFRWPTTPPRQRQDGE
EARRGEAIRKGRKSYLSMLNLQGISYVSGSLAPYQKLAADFDRKQWTRFP
EGSKTTPPHSHVDPRMKDYCPAVFRHLRLKLGVDPAEYMLAICGDTLRELASPKSG
SCFFTQDDRFMTKTVKKESEVKVLRMLRSYEHVROYKSLTLFRFGTHCKQKQCP
KIHRRFDLKGSSGRTIDTERKIDETTLQDLQVAFRLQRFVYELMKQIDMCP
FLETQGIMDYSLLGVMHFRNDYSVKIGISQIHAPKSGRKSGEGSSPFCGLCFVE
SGCCHRDILDSRKPRIQGINMPQADSESSKIIIDNPLNRHLPETPSSGSCCVYL
FPGIILIDYDITKLEHAYKSFVNDYIISADPKLYSRFPDPIRRVFLKEO"

16734..16878
/note="putative MITE, Galjin/Galjin-like"
19263..19454
/note="putative transposon, CACTA, En/Spm-like"
21819..24660
/gene="OJ1041F02.3"

CDS
complement(21819..22007,22096..22206,23772..23996,24523..24660)
/gene="OJ1041F02.3"

/codon_start=1
/product="hypothetical protein"
/protein_id="AAP06846.1"
/db_xref="GI:29893592"
/translation="WATSDSGSGSGEGRRLSIPKEGERIAPTRRPDGTLRKAIIRI
RAGYQSEVAVIYOSKGAQMRSGSDVPGYDPALDAKTKAAKRNRRRKQOAS
TTNDGKGLHIEDAGTNDPKMDVDSVDTISGAISESVVATSTDAATNSSES
SAPDIDKIRLAKKIRLAENQVQDPENLKPEQLEKKKIKGMEBELKLEKNSPA
AS"
complement(26563..30575)
/gene="OJ1041F02.4"

[illegible]

Db	19478	AGAACACCTTACTCCCAACATCACTGCTGATGCAC	119439
RESULT 11			
AC0844290/c			
LOCUS			
DEFINITION	AC0844290	55001 bp	DNA linear PRI 28-MAR-2002
ACCESSION	Homo sapiens	12 BAC RP11-56805 (Rosewell Park Cancer Institute Human	
VERSION	AC084290	complete sequence.	
KEYWORDS	AC084290.15	GI:19774277	
SOURCE ORGANISM	Homo sapiens (human)		
REFERENCE			
AUTHORS	Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C., Alsbrooks, S.L., Amaral-Lungue, H.C., Aye, R., Ayale, M., Banks, T., Barbieri, J., Benton, J., Bimsie, K., Blankensburg, K., Bonin, D., Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carrott, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chiu, D., Chowdhury, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dedrich, D.A., Delaney, K.R., Delgado, O., Dunn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Emertling, S., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Han, J., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hognes, M., Hollway, C., Hollins, B., Homisi, F., Howard, S., Huber, J., Hulik, S., Hume, J., Ioshikhes, I., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, B., Kelly, S., Khan, U., King, L., Korvath, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lee, E., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lien, C., Liu, J., Liu, W., Lonsheed, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapa, P., Marondei, I., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhinney, M., McLeod, M.P., Meador, M., Mei, G., Merscher, S., Metzner, M., Miller, A., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Montgomery, K.T., Morgan, M., Morris, S., Moser, M., Neal, D., Nelson, D., Newton, S., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okunolu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, Y., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojia, A., Rojiboken, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shin, C., Shooshtari, N., Sisom, I., Sodegriem, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabori, P., Tamerisa, A., Tamerisa, K., Tang, R., Tansley, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usumari, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wlezyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zortilla, S., Kuchelapatti, R., Weinstein, G., and Gibbs, R.		
TITLE	Unpublished		
JOURNAL	Direct Submission		
REFERENCE	2 (bases 1 to 55001)		
AUTHORS	Worley, K.C.		
TITLE	Direct Submission		
JOURNAL	Submitted (21-Oct-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
REFERENCE	3 (bases 1 to 55001)		
AUTHORS	Worley, K.C.		
TITLE	Direct Submission		
JOURNAL	Submitted (27-MAR-2002) Human Genome Sequencing Center, Department		

REFERENCE
AUTHORS
TITLE
JOURNAL

of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 55001)

Worley, K.C.

Direct Submission

Submitted (28-MAR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

On Mar 28, 2002 this sequence version replaced gi:19747063.

INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the
entire insert of this clone. Overlapping regions of clones are only
sequenced and submitted once, so the sequence for the remainder of
the insert may be found in the record for the adjacent clones.
Overlapping clones are noted at the beginning and end of the
Features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches
of a local database that includes entries from dbSTS, GDB, and
local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green,
unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST
(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
EST and cDNA sequences. Genes demonstrate at least two exons
flanked by consensus splice sites that maintained sequence
continuity across the splice junctions. Sequences that are not
identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum
standard of double strand coverage with a minimum of 2 clones and 2
reads with no ambiguities or 2 chemistries with a minimum of 2
clones and 3 reads with no ambiguities. If the sequence quality for
a region does not meet this standard, it will be indicated in the
annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality
standards - estimated error rate less than 1 per 10,000 bases.
Reports of lowest quality individual bases and measures of base
quality are listed below. Description of the metrics can be found
at URL:
<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT

FEATURES
source Location/Qualifiers

1. .55001
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="12"
/clone="RP11-568G5"
1. .1301
/note="overlaps bases 177088..178388 of clone AC079630"
/function="clone overlap"
35. .229
repeat_region
/rpt_family="MIR"
1002. .1023
/rpt_family="AT_rich"
1092. .1227
/rpt_family="MIR"
repeat_region
/rpt_family="L2"
complement(2156..2209)
repeat_region
/rpt_family="L2"
3563. .3650
/rpt_family="(CCGTA)n"
complement(3688..3765)
repeat_region
/rpt_family="L2"
complement(3840..3908)
repeat_region
/rpt_family="L1MD3"
complement(4234..4464)
repeat_region
/rpt_family="MIR"

repeat_region complement(4847..4914)
repeat_region /rpt_family="MIR"
4938..5053
/rpt_family="L2"
repeat_region 6350..6660
/rpt_family="AluSg"
repeat_region complement(7309..7743)
/rpt_family="LTR43"
7744..9144
/rpt_family="L1PA4"
complement(9151..9317)
repeat_region /rpt_family="LTR43"
10587..10846
/rpt_family="L1M4"
10912..11015
/rpt_family="L1P"
11096..11271
/rpt_family="L1ME"
11450..11635
/rpt_family="MLT11"
11941..12179
/rpt_family="MIR"
repeat_region 12797..12828
/rpt_family="(CATATA)n"
complement(12975..13087)
/rpt_family="FLAM_C"
13714..13737
/rpt_family="AT_rich"
14058..14219
/rpt_family="MER45"
14411..14545
/rpt_family="MIR"
repeat_region complement(15058..15127)
/rpt_family="MIR"
16186..16413
/rpt_family="MIR"
complement(17624..17922)
/rpt_family="AluDb"
complement(18309..18606)
/rpt_family="AluSx"
18844..18854
/rpt_family="AT_rich"
complement(18855..19135)
/rpt_family="AluDb"
19136..19158
/rpt_family="AT_rich"
complement(20660..20753)
/rpt_family="MER91C"
22179..22680
/rpt_family="(TAAA)n"
22180..22358
/rpt_family="AluJo"
22355..22632
/rpt_family="AluDb"
22633..22680
/rpt_family="(TAAA)n"
22682..22789
/rpt_family="AluJo"
22796..23099
/rpt_family="AluY"
complement(23147..23327)
/rpt_family="MIR"

Query Match

Best Local Similarity 60.0%; Score 24; DB 9; Length 55001;

Matches 27; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 ATACCTACTCCCAACATTCATCTTACTCAT 35
Db 47598 AATCTTACTCCCAACATTCATCTTACTCAT 47567

RESULT 12

AL807784/c
LOCUS AL807784 127196 bp DNA linear ROD 13-NOV-2002
DEFINITION Mouse DNA sequence from clone RP23-448C18 on chromosome X, complete
SEQUENCE.
ACCESSION AL807784
VERSION AL807784.11 GI:25045332
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS 1 (bases 1 to 127196)
TITLE Howden, P.
JOURNAL Direct Submission
Submitted (13-NOV-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone request: clonerequest@sanger.ac.uk
On Nov 15, 2002 this sequence version replaced gi:2493941.
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-448C18 is
from the RPCR-23 Mouse PAC library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBAC3.6.
Location/Qualifiers
source
1. 127196
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="X"
/clone="RP23-448C18"
/clone_1b="RPCI-23"

ORIGIN
Query Match 60.0%; Score 24; DB 10; Length 127196;
Best Local Similarity 75.0%; Pred. No. 43;
Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
1 AGGATACCTACTCCCAACATCATCTACTCATGCAC 40
Db 110413 AGGATACATAGCCAAATCACTTCTTACTAATGCCAC 110374

RESULT 13
AC148514/c
LOCUS AC148514 171940 bp DNA linear HTG 24-MAR-2004

DEFINITION Papio anubis clone RP41-375M17, WORKING DRAFT SEQUENCE. 3 ordered
pieces.
AC148514
AC148514.2 GI:45680463
VERSION HTG; HTGS PHASE2; HTGS DRAFT.
KEYWORDS Papio anubis (olive baboon)
SOURCE Papio anubis
ORGANISM Papio anubis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopitheciinae; Papio.
REFERENCE
AUTHORS 1 (bases 1 to 171940)
TITLE Antonellis, A., Ayele, K., Benjamin, B., Blakestey, R.W.,
Bouffard, G.G., Brinkley, C., Brooke, S., Chu, G., Coleman, B.,
Coleman, H., Daki, N., Engle, J., Granite, S., Guan, X., Gupta, J.,
Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Hu, P., Hurle, B.,
Idol, J.R., Jones, C., Karling, E., Kim, H., Kwong, P., Latic, P.,
Larsen, S., Lee-Lin, S.-O., Legaspi, R., Maduro, Q.L., Maduro, V.B.,
Margulis, E.H., Mastello, C., Maskeli, B., McDowell, J.,
Mullikin, J.C., Paguirigan, C., Portnoy, M.E., Prasad, A., Puri, O.,
Reddix-Dugue, N., Schandler, K., Schneider, M.G., Shah, K., Sison, C.,
Stantrop, S., Thomas, J.W., Thomas, P.J., Tsipouri, V., Vogt, J.L.,
Wetherby, K.D., Young, A. and Green, E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 171940)
REFERENCE
AUTHORS Green, E.D.
TITLE Direct Submission
Submitted (10-MAR-2004) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 171940)
REFERENCE
AUTHORS Green, E.D.
TITLE Direct Submission
Submitted (24-MAR-2004) NIH Intramural Sequencing Center, 8717
GroveMont Circle, Gaithersburg, MD 20877, USA
On Mar 24, 2004 this sequence version replaced gi:45332443.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc_zoo@hgtl.nih.gov
----- Project Information
Center project name: f5m
Center clone name: 375M17

The sequence data in this record represents an 'enhanced'
version of a Phase 2 submission. Specifically, the indicated
order and orientation of each sequence contig has been
established using one or more of the following: read-pair
data from individual subclones, overlaps with neighboring
clones, alignment with available reference sequence (e.g.,
human), and/or confirmation by PCR testing. In addition,
the sequence assembly is based on at least 8x average
coverage in Q20 bases and has been reviewed to rule out
gross misassemblies, the low-quality ends of sequence
contigs have been trimmed away, and each base is associated
with a Phrap-derived quality score.

----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 171395 bases at least Q40
Consensus quality: 171621 bases at least Q30
Consensus quality: 171701 bases at least Q20
Insert size: 173000; agarose-fp
Insert size: 171740; sum-of-contigs
Quality coverage: 8.27x in Q20 bases; agarose-fp
Quality coverage: 8.33x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the size
* of the gaps between them are based on estimates that have

```

* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
* 1 10148: contig of 10148 bp in length
* 10149 10248: gap of unknown length
* 10249 103023: contig of 92775 bp in length
* 103024 103123: gap of unknown length
* 103124 171940: contig of 68817 bp in length.
FEATURES
    source
        1..171940
            /organism="Papio anubis"
            /mol_type="genomic DNA"
            /db_xref="taxon:9555"
            /clone="RP41-375M17"
            /clone_1lb="RP41"
            /note="BAC resource: http://bacpac.chori.org/"
        1..10148
            /note="assembly_fragment"
            clone_end:SP6
            vector_side:left"
        10249..103023
            /note="assembly_fragment"
            103124..171940
            /note="assembly_fragment"
            clone_end:T7
            vector_side:right"
ORIGIN
Query Match      60.0%; Score 24; DB 2; Length 171940;
Best Local Similarity 84.4%; Pred. NO. 42;
Matches 27; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
OY      4 ATACCTACTCCCAACATCCATCTACTCAT 35
Db      33556 AAACCTACTCCCAACATCTACTCAT 33525
RESULT 14
AC148501/c
LOCUS
DEFINITION
AC148501      193735 bp      DNA      linear      HTG 09-JUL-2004
Callithrix jacchus clone CH259-368E20, WORKING DRAFT SEQUENCE, 11
ordered pieces.
AC148501
AC148501.3 GI:50080339
HTG; HTGS PHASE2; HTGS DRAFT.
SOURCE
Callithrix jacchus (white-tufted-ear marmoset)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae;
Callitrix.
1 (bases 1 to 193735)
Antorellis A., Ayele, K., Benjamin, B., Blakeley, R.W.,
Bouffard, G.G., Brinkley, C., Brooks, S., Chu, G., Coleman, B.,
Coleman, H., Daki, N., Engle, D., Guan, X., Gupta, J., Haghighi, P.,
Han, J., Hansen, N., Ho, S.-L., Hu, P., Hurle, B., Idol, J.R., Jones, C.,
Karling, E., Kim, H., Kwong, P., Latic, P., Larson, S., Lee-Lin, S.-O.,
Legaspi, R., Madden, M., Maduro, O.L., Maduro, V.B., Margulies, E.H.,
Masello, C., Maskeri, B., Mcowell, J., Mullikin, J.C., Peguiri, C.,
Park, M., Portnoy, M.E., Prasad, A., Puri, O., Reddix-Dugue, N.,
Schandler, K., Schueler, M.G., Shah, K., Sison, C., Stantirip, S.,
Thomas, J.W., Thomas, P.J., Tsipouri, V., Vogt, J.L., Wetherby, K.D.,
Young, A. and Green, E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 193735)
Green, E.D.
Direct Submission
JOURNAL
Submitted (10-MAR-2004) NIH Intramural Sequencing Center, 8717
Groveom Circle, Galthersburg, MD 20877, USA
3 (bases 1 to 193735)
Green, E.D.
Direct Submission

```

```

JOURNAL
Submitted (09-JUL-2004) NIH Intramural Sequencing Center, 8717
Groveom Circle, Galthersburg, MD 20877, USA
On Jul 9, 2004 this sequence version replaced gi:46430810.
-----
Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoo@nih.gov
-----
Project Information
Center project name: ffw
Center clone name: 368E20

The sequence data in this record represents an 'enhanced'
version of a phase 2 submission. Specifically, the indicated
order and orientation of each sequence contig has been
established using one or more of the following: read-pair
data from individual subclones, overlaps with neighboring
clones, alignment with available reference sequence (e.g.,
human), and/or confirmation by PCR testing. In addition,
the sequence assembly is based on at least 8x average
coverage in Q20 bases and has been reviewed to rule out
gross misassemblies, the low-quality ends of sequence
contigs have been trimmed away, and each base is associated
with a Phrap-derived quality score.
-----
Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990119
Consensus quality: 190717 bases at least Q40
Consensus quality: 191709 bases at least Q30
Consensus quality: 193323 bases at least Q20
Insert size: 14400; agarose-fp
Insert size: 192735; sum-of-contigs
Quality coverage: 13.33x in Q20 bases; agarose-fp
Quality coverage: 9.96x in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 11 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
1 29309: contig of 29309 bp in length
29310 29409: gap of unknown length
29410 37712: contig of 8303 bp in length
37713 37812: gap of unknown length
37813 40472: contig of 2660 bp in length
40473 40572: gap of unknown length
40573 45724: contig of 5152 bp in length
45725 45824: gap of unknown length
45825 47844: gap of 2020 bp in length
47845 47944: gap of unknown length
47945 61027: contig of 13083 bp in length
61028 61127: gap of unknown length
61128 63312: contig of 2085 bp in length
63313 70560: gap of unknown length
70561 70660: gap of 7248 bp in length
70661 73658: contig of 2998 bp in length
73659 73758: gap of unknown length
73759 103653: contig of 29895 bp in length
103654 103753: gap of unknown length
103754 193735: contig of 89982 bp in length.
-----
Location/Qualifiers
1..193735
/organism="Callithrix jacchus"
/mol_type="genomic DNA"
/db_xref="taxon:9483"
/clone="CH259-368E20"
/clone_1lb="CH259"
FEATURES
    source

```

```
misc_feature      /note="BAC resource: http://bacpac.chori.org/"
1..29309
/note="assembly_fragment
clone_end:17
vector_side:left"
29410..37712
/note="assembly_fragment"
37813..40472
/note="assembly_fragment"
40573..45724
/note="assembly_fragment"
45825..47844
/note="assembly_fragment"
47945..61027
/note="assembly_fragment"
61128..63212
/note="assembly_fragment"
63313..70560
/note="assembly_fragment"
70661..73658
/note="assembly_fragment"
73759..103653
/note="assembly_fragment"
103754..193735
/note="assembly_fragment
clone_end:SP6
vector_side:right"
127712..193735
/note="clone overlaps with GenBank Accession Number
AC148551 clone CH259-256024 (center project name fvv)"
misc_feature
misc_feature
misc_feature
misc_feature
misc_feature
misc_feature
misc_feature
misc_feature
misc_feature
misc_feature
```

ORIGIN

```
Query Match      59.0%; Score 23.6; DB 2; Length 193735;
Best Local Similarity 76.3%; Pred. No. 60;
Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
```

```
QY      2  GGATACCTACTCCCAACATGCATCTTACTCATGCAA 39
Db      89109  GGAAATCACTCCAAAATCCCTCAAAATCATGCA 89072
```

```
RESULT 15
AC008897      117026 bp      DNA      linear      PRI 07-NOV-2000
DEFINITION    Homo sapiens chromosome 5 clone CTD-2235C13, complete sequence.
AC008897
VERSION       AC008897.7 GI:11139448
KEYWORDS
SOURCE
ORGANISM      Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 117026)
DOE Joint Genome Institute and Stanford Human Genome Center.
Direct Submision
Unpublished
2 (bases 1 to 117026)
DOE Joint Genome Institute.
Direct Submision
Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 117026)
DOE Joint Genome Institute and Stanford Human Genome Center.
Direct Submision
Submitted (07-NOV-2000) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
On Nov 7, 2000 this sequence version replaced gi:7713370.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.5% of Sequence;
Estimated Total Number of Errors is 0.5.
```

```
STS Content:
WI-9815 G05428
SHGC-16751 G15410
WI-7171 G06431.
FEATURES
Source      Location/Qualifiers
1..117026
```

```
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTD-2235C13"
```

ORIGIN

```
Query Match      57.5%; Score 23; DB 9; Length 117026;
Best Local Similarity 74.4%; Pred. No. 1.1e+02;
Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
```

```
QY      1  AGGATACCTACTCCCAACATGCATCTTACTCATGCAA 39
Db      72634  AGGAACGGATTACAACTATCCACCTTTCTCAAGCA 72672
```

```
Search completed: September 16, 2005, 03:01:29
Job time : 139.222 secs
```

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using bw model

Run on: September 15, 2005, 08:10:17 ; Search time 19.3656 Seconds
(without alignments)
12227.321 Million cell updates/sec

Title: US-10-713-381-1_COPY_1239_1278

Perfect score: 40

Sequence: 1 aggatactaccctcccaacacatccatctactcgtcgaac 40

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : N_GeneSeq_16Dec04:*
1: geneSeqn19808:*
2: geneSeqn19909:*
3: geneSeqn20008:*
4: geneSeqn20018:*
5: geneSeqn20028:*
6: geneSeqn20038:*
7: geneSeqn20048:*
8: geneSeqn20058:*
9: geneSeqn20068:*
10: geneSeqn20078:*
11: geneSeqn20088:*
12: geneSeqn20098:*
13: geneSeqn20108:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	40	100.0	40	5	AAH76337	Aah76337 Z. mays M
2	40	100.0	50	5	AAH76336	Aah76336 Z. mays M
3	40	100.0	158	5	AAH76334	Aah76334 Z. mays M
4	40	100.0	255	5	AAH76340	Aah76340 Z. mays M
5	40	100.0	1394	2	AAH07408	Aax07408 Zea mays
6	40	100.0	1394	2	AAH07409	Aax07409 Zea mays
7	40	100.0	1394	5	AAH76332	Aah76332 Z. mays M
8	40	100.0	1394	5	AAH76333	Aah76333 Z. mays M
9	22.4	56.0	226215	11	ACN45146	Acn45146 Human gene
10	21.6	54.0	4110	8	ABZ10202	Abz10202 Haematopo
11	21.6	54.0	6309	6	ABL23204	AbL23204 Human imm
12	21.6	54.0	7110	10	ADB54282	AdB54282 Precreate
13	21.6	54.0	7110	10	ADB54282	AdB54282 Human lym
14	21.6	54.0	7110	13	ADG89580	AdG89580 Oligonuc
15	21.6	53.5	1341	6	ABZ13750	Abz13750 Arabidops
16	21.4	53.5	6681	6	ABL32155	AbL32155 Human imm
17	21.4	53.5	6681	6	ABL54304	AbL54304 Chemical
18	21.4	53.5	8168	6	ABL2928	AbL2928 Human imm
19	21.4	53.5	8168	6	AAH53328	AaH53328 Chemical
20	21.4	53.5	24401	4	ABL03396	AbL03396 Drosophi

21	21	52.5	2000	12	ADJ41594	Adj41594 Plant CDN
22	21	52.5	2792	12	ADM98941	Adm98941 Diterpene
23	21	52.5	3117	12	ADM98942	Adm98942 Diterpene
24	21	52.5	5507	8	ABZ10204	Abz10204 Haematopo
25	21	52.5	5507	8	ABZ10058	Abz10058 Haematopo
26	21	52.5	8404	4	AAH46500	AaH46500 Tumour su
27	21	52.5	8404	6	ABL33595	AbL33595 Human imm
28	21	52.5	8404	10	ADB54216	AdB54216 Precreate
29	21	52.5	8404	10	ADB54088	AdB54088 Precreate
30	21	52.5	8404	10	ADB84076	AdB84076 Human lym
31	21	52.5	8404	10	ADB84152	AdB84152 Human lym
32	21	52.5	8404	13	ADG89242	AdG89242 Oligonuc
33	21	52.5	8404	13	ADG89516	AdG89516 Oligonuc
34	21	52.5	10286	4	AAH53308	AaH53308 Chemicall
35	21	52.5	10286	6	ABK28147	AbK28147 DNA trans
36	21	52.5	63115	10	ADC85419	AdC85419 Mouse Tle
37	21	52.5	63294	9	ADA02939	Ada02939 Mouse Tle
38	21	52.5	63294	10	ADM72677	AdM72677 Mouse Tle
39	21	52.5	63294	12	ADM74534	AdM74534 Murine ca
40	21	52.5	106315	11	ACN43966	AcN43966 Human gen
41	20.8	52.0	891	8	ABZ51893	Abz51893 Aspergill
42	20.8	52.0	6620	4	AAH54548	AaH54548 Chemicall
43	20.8	52.0	6620	6	ABK28416	AbK28416 DNA trans
44	20.8	52.0	6620	6	ABN80279	AbN80279 Human che
45	20.8	52.0	110000	5	AAH61373_1	Continuation (2 of

ALIGNMENTS

RESULT 1	
AAH76337	AAH76337 standard; DNA; 40 BP.
XX	
AC	AAH76337;
XX	
DT	29-OCT-2001 (first entry)
XX	
DE	Z. mays Ms45 male tissue-preferred regulatory region fragment.
XX	
KM	Ms45; male tissue; regulatory region; transcription; male fertility;
XX	
KW	hybrid seed; de.
OS	Zea mays.
XX	
PN	WO200160997-A2.
XX	
PD	23-AUG-2001.
XX	
PF	13-FEB-2001; 2001WO-US004527.
XX	
PR	15-FEB-2000; 2000US-00504487.
XX	
PA	(PION-) PIONEER HI-BRED INT INC.
XX	
PI	Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX	
DR	WPI; 2001-514772/56.
XX	
PT	A male tissue-preferred regulatory region comprising nucleotide sequences
XX	
PT	essential for initiating transcription of the Ms45 gene useful for
XX	
PT	mediating fertility in a male plant.
XX	
PS	Claim 14; Page 32; 50pp; English.
XX	
CC	The invention provides a male tissue-preferred regulatory region (I)
XX	
CC	comprising nucleotide sequences essential for initiating transcription of
XX	
CC	the Ms45 gene. A method of mediating male fertility in a plant is
XX	
CC	provided that involves introducing an expression vector comprising a
XX	
CC	promoter operably linked to (I) into a plant where the exogenous gene
XX	
CC	impacts male fertility of the plant and (II) controls expression of the
XX	
CC	exogenous gene. A method of producing hybrid seeds is also provided. The
XX	
CC	present sequence represents a DNA fragment upstream of the TATA box of a

```
CC 2. mays Ms45 male-tissue preferred regulatory region nucleotide sequence
XX
SQ Sequence 40 BP; 14 A; 14 C; 3 G; 9 T; 0 U; 0 Other;
Query Match 100.0%; Score 40; DB 5; Length 40;
Best Local Similarity 100.0%; Pred. No. 3.6e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGATACCTACTCTCCCAACATCCATCTTACTATGCAAC 40
DB 1 AGGATACCTACTCTCCCAACATCCATCTTACTATGCAAC 40

RESULT 2
AAH76336
ID AAH76336 standard; DNA; 50 BP.
AC AAH76336;
XX
XX
DT 29-OCT-2001 (first entry)
XX
DE 2. mays Ms45 male tissue-preferred regulatory region fragment.
XX
XX Ms45; male tissue; regulatory region; transcription; male fertility;
XX hybrid seed; ds.
XX
OS Zea mays.
XX
XX WO200160997-A2.
XX
XX 23-AUG-2001.
XX
XX 13-FEB-2001; 2001WO-US004527.
XX
XX 15-FEB-2000; 2000US-00504487.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX WPI; 2001-514772/56.
XX
XX A male tissue-preferred regulatory region comprising nucleotide sequences
XX essential for initiating transcription of the Ms45 gene useful for
XX mediating fertility in a male plant.
XX
XX Claim 5; Page 47; 50pp; English.
XX
XX The invention provides a male tissue-preferred regulatory region (I)
XX comprising nucleotide sequences essential for initiating transcription of
XX the Ms45 gene. A method of mediating male fertility in a plant is
XX provided that involves introducing an expression vector comprising a
XX promoter operably linked to (I) into a plant where the exogenous gene
XX impacts male fertility of the plant and (I) controls expression of the
XX exogenous gene. A method of producing hybrid seeds is also provided. The
XX present sequence represents a DNA fragment -72 to -111 bases upstream of
XX the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
XX nucleotide sequence
XX
SQ Sequence 50 BP; 18 A; 17 C; 5 G; 10 T; 0 U; 0 Other;
Query Match 100.0%; Score 40; DB 5; Length 50;
Best Local Similarity 100.0%; Pred. No. 3.7e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGATACCTACTCTCCCAACATCCATCTTACTATGCAAC 40
DB 11 AGGATACCTACTCTCCCAACATCCATCTTACTATGCAAC 50

RESULT 3
AAH76334
ID AAH76334 standard; DNA; 158 BP.
```

```
XX
XX AAH76334;
XX
XX 29-OCT-2001 (first entry)
XX
XX 2. mays Ms45 male tissue-preferred regulatory region fragment.
XX
XX Ms45; male tissue; regulatory region; transcription; male fertility;
XX hybrid seed; ds.
XX
XX Zea mays.
XX
XX WO200160997-A2.
XX
XX 23-AUG-2001.
XX
XX 13-FEB-2001; 2001WO-US004527.
XX
XX 15-FEB-2000; 2000US-00504487.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX WPI; 2001-514772/56.
XX
XX A male tissue-preferred regulatory region comprising nucleotide sequences
XX essential for initiating transcription of the Ms45 gene useful for
XX mediating fertility in a male plant.
XX
XX Claim 5; Page 47; 50pp; English.
XX
XX The invention provides a male tissue-preferred regulatory region (I)
XX comprising nucleotide sequences essential for initiating transcription of
XX the Ms45 gene. A method of mediating male fertility in a plant is
XX provided that involves introducing an expression vector comprising a
XX promoter operably linked to (I) into a plant where the exogenous gene
XX impacts male fertility of the plant and (I) controls expression of the
XX exogenous gene. A method of producing hybrid seeds is also provided. The
XX present sequence represents a DNA fragment -38 to -195 bases upstream of
XX the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
XX nucleotide sequence
XX
SQ Sequence 158 BP; 41 A; 50 C; 21 G; 46 T; 0 U; 0 Other;
Query Match 100.0%; Score 40; DB 5; Length 158;
Best Local Similarity 100.0%; Pred. No. 4.4e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGATACCTACTCTCCCAACATCCATCTTACTATGCAAC 40
DB 86 AGGATACCTACTCTCCCAACATCCATCTTACTATGCAAC 125

RESULT 4
AAH76340
ID AAH76340 standard; DNA; 255 BP.
AC AAH76340;
XX
XX 29-OCT-2001 (first entry)
XX
XX 2. mays Ms45 promoter fragment.
XX
XX Ms45; male tissue; regulatory region; transcription; male fertility;
XX hybrid seed; promoter; ds.
XX
XX Zea mays.
XX
XX WO200160997-A2.
XX
XX 23-AUG-2001.
```

PF 13-FEB-2001; 2001WO-US004527.
 XX 15-FEB-2000; 2000US-00504487.
 XX (PION-) PIONEER HI-BRED INT INC.
 XX
 XX PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
 XX WPI; 2001-514772/56.
 XX
 PT A male tissue-preferred regulatory region comprising nucleotide sequences
 PT essential for initiating transcription of the Ms45 gene useful for
 PT mediating fertility in a male plant.
 XX
 PS Example 5; Fig 8; 50pp; English.
 XX
 CC The invention provides a male tissue-preferred regulatory region (I)
 CC comprising nucleotide sequences essential for initiating transcription of
 CC the Ms45 gene. A method of mediating male fertility in a plant is
 CC provided that involves introducing an expression vector comprising a
 CC promoter operably linked to (I) into a plant where the exogenous gene
 CC impacts male fertility of the plant and (I) controls expression of the
 CC exogenous gene. A method of producing hybrid seeds is also provided. The
 CC present sequence represents a Z. mays Ms45 promoter fragment
 XX
 SQ Sequence 255 BP; 59 A; 86 C; 39 G; 71 T; 0 U; 0 Other;
 Query Match 100.0%; Score 40; DB 5; Length 255;
 Best Local Similarity 100.0%; Pred. No. 4.7e-06;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AGGATACCTTACTCTCCCAACATCATCTTACTCATGCAAC 40
 DB 99 AGGATACCTTACTCTCCCAACATCATCTTACTCATGCAAC 138
 XX
 RESULT 5
 AAX07408
 ID AAX07408 standard; DNA; 1394 BP.
 XX
 AC AAX07408;
 XX
 DT 08-JUN-1999 (first entry)
 XX
 DE Zea mays Ms45 male tissue-preferred regulatory region.
 XX
 KM Me45; male; tissue-preferred; regulatory region; plant cells;
 KM plant tissue; differentiated; maize; hybrid seed; fertility; ss.
 XX
 OS Zea mays.
 XX
 PN WO9859061-A1.
 XX
 PD 30-DEC-1998.
 XX
 PF 19-JUN-1998; 98WO-US012895.
 XX
 PR 23-JUN-1997; 97US-00880499.
 XX
 PA (PION-) PIONEER HI-BRED INT INC.
 XX
 PI Albertsen MC, Fox TW, Garnaat CW, Huffman GA, Kendall TL;
 XX WPI; 1999-105628/09.
 XX
 PT New nucleic acid encoding a Ms45 male tissue-preferred regulatory region
 PT - useful in mediating plant fertility, especially hybrid seed production.
 XX
 PS Claim 2; Page 22-23; 39pp; English.
 XX
 CC The sequence is that encoding an Ms45 male tissue-preferred regulatory
 CC region. It may be used in the construction of a vector for a method of
 CC producing exogenous genes in a male tissue- preferred manner, which is

CC useful in restoring or conferring fertility, such as in hybrid seed
 CC production. In conferring fertility, a monocot/dicot plant is transformed
 CC with the exogenous nucleotide sequence (a male sterility gene, preferably
 CC Ms45), which encodes a product selected from auxins, rolB and diptheria
 CC toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
 CC and infertile plants
 XX
 SQ Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;
 Query Match 100.0%; Score 40; DB 2; Length 1394;
 Best Local Similarity 100.0%; Pred. No. 6.1e-06;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AGGATACCTTACTCTCCCAACATCATCTTACTCATGCAAC 40
 DB 1239 AGGATACCTTACTCTCCCAACATCATCTTACTCATGCAAC 1278
 XX
 RESULT 6
 AAX07409
 ID AAX07409 standard; DNA; 1394 BP.
 XX
 AC AAX07409;
 XX
 DT 08-JUN-1999 (first entry)
 XX
 DE Zea mays Ms45 male tissue-preferred regulatory region.
 XX
 KM Me45; male; tissue-preferred; regulatory region; plant cells;
 KM plant tissue; differentiated; hybrid seed; fertility; ss.
 XX
 OS Zea mays.
 XX
 PN WO9859061-A1.
 XX
 PD 30-DEC-1998.
 XX
 PF 19-JUN-1998; 98WO-US012895.
 XX
 PR 23-JUN-1997; 97US-00880499.
 XX
 PA (PION-) PIONEER HI-BRED INT INC.
 XX
 PI Albertsen MC, Fox TW, Garnaat CW, Huffman GA, Kendall TL;
 XX WPI; 1999-105628/09.
 XX
 PT New nucleic acid encoding a Ms45 male tissue-preferred regulatory region
 PT - useful in mediating plant fertility, especially hybrid seed production.
 XX
 PS Claim 3; Page 23-24; 39pp; English.
 XX
 CC The sequence is that encoding an Ms45 male tissue-preferred regulatory
 CC region. It may be used in the construction of a vector for a method of
 CC producing exogenous genes in a male tissue- preferred manner, which is
 CC useful in restoring or conferring fertility, such as in hybrid seed
 CC production. In conferring fertility, a monocot/dicot plant is transformed
 CC with the exogenous nucleotide sequence (a male sterility gene, preferably
 CC Ms45), which encodes a product selected from auxins, rolB and diptheria
 CC toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
 CC and infertile plants
 XX
 SQ Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;
 Query Match 100.0%; Score 40; DB 2; Length 1394;
 Best Local Similarity 100.0%; Pred. No. 6.1e-06;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AGGATACCTTACTCTCCCAACATCATCTTACTCATGCAAC 40
 DB 1239 AGGATACCTTACTCTCCCAACATCATCTTACTCATGCAAC 1278

```
RESULT 7
AAH76332
ID AAH76332 standard; DNA; 1394 BP.
XX
AC AAH76332;
XX
DT 29-OCT-2001 (first entry)
XX
DE Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.
XX
KW Ms45; male tissue; regulatory region; transcription; male fertility;
XX hybrid seed; ds.
XX
OS Zea mays.
XX
PN WO200160997-A2.
XX
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
DR WPI; 2001-514772/56.
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the Ms45 gene useful for
PT mediating fertility in a male plant.
XX
PS Claim 4; Page 46; 50pp; English.
XX
CC The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the Ms45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a nucleic acid sequence encoding an Ms45 male
CC -tissue preferred regulatory region from Z. mays
XX
SQ Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 40; DB 5; Length 1394;
Best Local Similarity 100.0%; Pred. No. 6.1e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGATACCTACTCCCAACAATCCATCTTACTGATGCAAC 40
DB 1239 AGGATACCTACTCCCAACAATCCATCTTACTGATGCAAC 1278
XX
RESULT 8
AAH76333
ID AAH76333 standard; DNA; 1394 BP.
XX
AC AAH76333;
XX
DT 29-OCT-2001 (first entry)
XX
DE Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.
XX
KW Ms45; male tissue; regulatory region; transcription; male fertility;
XX hybrid seed; ds.
XX
OS Zea mays.
XX
PN WO200160997-A2.
XX
```

```
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
DR WPI; 2001-514772/56.
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the Ms45 gene useful for
PT mediating fertility in a male plant.
XX
PS Claim 4; Page 47; 50pp; English.
XX
CC The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the Ms45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a nucleic acid sequence encoding an Ms45 male
CC -tissue preferred regulatory region from Z. mays
XX
SQ Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 40; DB 5; Length 1394;
Best Local Similarity 100.0%; Pred. No. 6.1e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGATACCTACTCCCAACAATCCATCTTACTGATGCAAC 40
DB 1239 AGGATACCTACTCCCAACAATCCATCTTACTGATGCAAC 1278
XX
RESULT 9
ACN45146
ID ACN45146 standard; DNA; 226215 BP.
XX
AC ACN45146;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human genomic sequence hCG1639824.
XX
KW Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003073826-A2.
XX
PD 12-SEP-2003.
XX
PF 28-FEB-2003; 2003WO-US006235.
XX
PR 01-MAR-2002; 2002US-00087192.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW;
XX
DR WPI; 2003-328604/31.
XX
PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
PT comprises a nucleotide sequence.
XX
PS Claim 1; SEQ ID NO 1948; opp; English.
XX
CC The present invention relates to novel DNA and protein sequences which
```

are associated with carcinomas. The sequences are useful for: (i) for screening drug candidates; (ii) for screening of bioactive agent capable of binding to CarcinoMa Associated Protein (CAP); (iii) for screening of a bioactive agent capable of modulating the activity of CAP; (iv) for evaluating the effect of a candidate carcinoma drug; (v) for diagnosing carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a bioclip; (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for determining CarcinoMa Associated (CA) gene copy number. In addition, the CA genes are useful as DNA vaccines and the CAP are useful as markers of carcinoma including lymphoma. The present sequence is one such CA coding sequence. Note: This patent is an equivalent to basic patent US2002182586A1, for which no sequence data was published

Sequence 226215 BP; 65331 A; 42804 C; 45218 G; 72862 T; 0 U; 0 Other;

Query Match 56.0%; Score 22.4; DB 11; Length 226215;
Best Local Similarity 72.5%; Pred. No. 1.2e+02;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

1 AGGATACCTACTCCCAACATCCATCTTACTCATGCAAC 40
183206 AGAATTCCTGCTTCCAAACACAGCATCTTACTCTTTAC 183245

RESULT 10
AB210202/C
ID AB210202 standard; DNA; 4110 BP.
XX
AC AB210202;
XX
DT 16-JUN-2003 (first entry)
XX
DE Haematopoietic cell proliferation disorder related DNA sequence #342.
XX
KM Human; haematopoietic cell proliferation disorder; cytostatic;
KM Gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
KM Cytosine methylation state; gene; db.
OS Homo sapiens.
XX
PN MO200277272-A2.
XX
PD 03-OCT-2002.
XX
PF 26-MAR-2002; 2002MO-EP003401.
XX
PR 26-MAR-2001; 2001US-0278333P.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Berlin K, Braun A, Distler J, Guecig D, Howe A, Mueller J;
PI Olek A, Piepenbrock C, Adorjan P, Grabs G, Lesche R, Leu B;
PI Lewin A, Lipscher E, Mater S, Model F, Mueller V, Otto T, Pellet C;
PI Schwabe I, Ziebarth H;
XX
DR WPI, 2003-018942/01.
XX
PT Detecting and differentiating between hematopoietic cell proliferative
PT disorders, comprises connecting a target nucleic acid with a reagent that
PT distinguishes between methylated and non-methylated Cpg dinucleotides.
XX
XX Claim 28; SEQ ID NO 342; 117bp; English.
XX
XX The present invention describes a method for detecting and
XX differentiating between haematopoietic cell proliferative disorders
XX associated with at least 1 gene and/or their regulatory regions in a
XX subject. The method comprises contacting a target nucleic acid in a
XX biological sample obtained from the subject with at least 1 reagent,
XX which distinguishes between methylated and non-methylated Cpg
XX dinucleotides within the target nucleic acid. AB209861 to AB211118
XX represent specifically claimed nucleotide sequences from the present
XX invention. Oligonucleotides from the present invention can be used: for

differentiating between healthy haematopoietic cells and proliferative
disorder haematopoietic cells; for differentiating between acute
lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
determining the cytosine methylation state and/or single nucleotide
polymorphisms (SNPs) of haematopoietic cell proliferation disorder
related sequences and their complements; and as primers for the
amplification of haematopoietic cell proliferation disorder related DNA
sequences. The nucleotide sequences from the present invention can also
be used for detecting a predisposition to, differentiation between
subclasses, diagnosis, prognosis, treatment and/or monitoring of
haematopoietic cell proliferation disorders. The present method enables a
highly specific classification of haematopoietic cell proliferative
disorders allowing for improved and informed treatment of patients

Sequence 4110 BP; 945 A; 0 C; 1132 G; 2033 T; 0 U; 0 Other;

Query Match 54.0%; Score 21.6; DB 8; Length 4110;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

4 ATACCTACTCTCCCAACATCCATCTTAC 31
3083 AACCTAACCCCAACATCCATCTTAC 3056

RESULT 11
ABL32304/C
ID ABL32304 standard; DNA; 6309 BP.
XX
AC ABL32304;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human immune system associated gene SEQ ID NO: 277.
XX
KM Human; immune system associated; cytosine methylation; antiasthmatic;
KM antitreticlerotic; antrianemic; cyrostatic; noctropic;
KM neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KM antiinflammatory; antiarthritic; antidiabetic; antipsoriatic;
KM acute myeloid leukaemia; cancer; eye disease; arteriosclerosis; anaemia;
KM neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
KM db.
XX
OS Homo sapiens.
XX
PN MO200200928-A2.
XX
PD 03-JUN-2002.
XX
PF 02-JUL-2001; 2001WO-EP007537.
XX
PR 30-JUN-2000; 2000DE-01032529.
XX
PR 01-SEP-2000; 2000DE-01043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI, 2002-130909/17.
XX
PT Nucleic acid comprising fragment of chemically modified gene, useful for
PT diagnosis and treatment of diseases associated with abnormal cytosine
PT methylation.
XX
XX Claim 1; SEQ ID NO 277; 32bp + Sequence Listing; German.
XX
XX The present invention provides a number of human immune system associated
XX genes which are modified by the methylation of cytosines. The sequences
XX can be used in the diagnosis and treatment of immune system disorders,
XX including eye diseases such as retinopathy, neovascular glaucoma and
XX macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
XX leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,

CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention
XX
SO Sequence 6309 BP; 1758 A; 69 C; 1364 G; 3118 T; 0 U; 0 Other;

Query Match 54.0%; Score 21.6; DB 6; Length 6309;
Best Local Similarity 75.0%; Pred. No. 1.5e+02;
Matches 27; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 4 ATACCTACTCCCAAGATCCATCTTACATGCAA 39
DB 3335 ATACCTATTCCCATGACCAATCTCATATATCCAA 3300

RESULT 12
ADB54282/C
ID ADB54282 standard; DNA; 7110 BP.

XX ADB54282;

DT 04-DEC-2003 (first entry)

DE Pretreated genomic DNA region 206.

KW colon cell proliferative disorder; non methylated CpG dinucleotide;
KM cytostatic; cancer; adenoma; carcinoma; cytosine methylation state; ds.

OS Unidentified.

XX WO2003072821-A2.

PD 04-SEP-2003.

PF 27-FEB-2003; 2003WO-EP002035.

PR 27-FEB-2002; 2002EP-00004551.

PA (EPIG-) EPIGENOMICS AG.

PI Adorjan P, Burger M, Maier S, Nimmrich I, Becker E, Lesche R;
PI Rujan T, Schmitt A;

DR WPI; 2003-731620/69.

PT Detecting and differentiating between colon cell proliferative disorders
PT associated with a gene or its regulatory regions comprises contacting a
PT target nucleic acid in a biological sample obtained from the subject with
PT a reagent.

PS Claim 32; SEQ ID NO 338; 74bp; English.

CC The invention relates to a novel method for detecting and differentiating
CC between colon cell proliferative disorders associated with at least one
CC gene or its regulatory regions. The method comprises contacting a target
CC nucleic acid in a biological sample obtained from the subject with at
CC least one reagent or a series of reagents, where the reagent or series of
CC reagents, distinguishes between methylated and non methylated CpG
CC dinucleotides within the target nucleic acid. The molecules of the
CC invention demonstrate cytosstatic activity whilst the method may useful
CC for detecting and differentiating between colon cell proliferative
CC disorders, including cancers such as colon adenoma and colon carcinoma.
CC The PNA (peptide nucleic acid)-oligomers are useful as probes for
CC determining cytosine methylation state or single nucleotide
CC polymorphisms. The current sequence is that of the pretreated genomic DNA
CC region of the invention. This sequence is not shown within the
CC specification but is taken from Wipoweb.

XX Sequence 7110 BP; 1818 A; 0 C; 1721 G; 3571 T; 0 U; 0 Other;

Query Match 54.0%; Score 21.6; DB 10; Length 7110;
Best Local Similarity 85.7%; Pred. No. 1.5e+02;
Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 ATACCTACTCCCAAGATCCATCTTAC 31
DB 3083 AAACCTAACCCAAACATCTCCTAC 3056

RESULT 13
ADE84196/C
ID ADE84196 standard; DNA; 7110 BP.

XX ADE84196;

DT 29-JAN-2004 (first entry)

DE Human lymphoid cell proliferative disorder gene derived DNA #132.

KW ds; lymphoid cell proliferative disorder; methylation;
KM methylated CpG dinucleotide; single nucleotide polymorphism; SNP;
KM diffuse large B-cell lymphoma; mantle cell lymphoma;
KM chronic lymphocytic leukemia; small lymphocytic lymphoma;
KM follicular lymphoma; diagnosis; prognosis.

OS Homo sapiens.

XX WO2003044226-A2.

PD 30-MAY-2003.

PF 25-NOV-2002; 2002WO-EP033265.

PR 23-NOV-2001; 2001DE-01057491.

PR 28-DEC-2001; 2001DE-01064501.

PA (EPIG-) EPIGENOMICS AG.

PI Burger M, Caldwell C, Genc B, Becker E, Maier S, Nimmrich I;
PI WPI; 2003-457621/43.

PT Detecting and differentiating between lymphoid cell proliferative
PT disorders comprises contacting a target nucleic acid with at least one
PT reagent that distinguishes between methylated and non-methylated CpG
PT dinucleotides.

PS Claim 26; SEQ ID NO 192; 448bp; English.

CC The invention relates to a method of detecting and differentiating
CC between lymphoid cell proliferative disorders associated with at least
CC one gene and/or their regulatory regions in a subject by contacting a
CC target nucleic acid in a biological sample obtained from the subject with
CC at least one reagent or series of reagents that distinguish between
CC methylated and non-methylated CpG dinucleotides within the target nucleic
CC acid. The genes and/or their regulatory regions are preferably selected
CC from MRL1, CSNK2B, EGR4, AR, CDK4, RB2, CDC25A, GRB beta, MYO1, CDH3,
CC MYC1, ELK1, ABL1, APC, BC12, CDH1, CDKN1A, CDKN1B, CDKN2A, FOS,
CC GSTR1, HIC-1, MGMT, MLH1, MOS, MYC, PTEN, RBL2, TGFBR2, TP73, CDKN1C,
CC GSK3beta, ESR1, APT1, BAK1, BAX or HOKA5. Oligomers, peptide nucleic
CC acid (PNA)-oligomers and/or isolated nucleic acids based on the sequences
CC of the genes are useful for detecting the methylation state of all the
CC CpG dinucleotides within one or more the sequences, or their complements,
CC for determining the cytosine methylation state and or single nucleotide
CC polymorphisms (SNPs), and for differentiating at least two of the medical
CC conditions such as diffuse large B-cell lymphoma, mantle cell lymphoma,
CC chronic lymphocytic leukemia, small lymphocytic lymphoma and follicular
CC lymphoma. They are also useful for detecting of a predisposition to,
CC differentiation between subclasses, diagnosis, prognosis, treating and/or
CC monitoring of lymphoid cell proliferative disorder. This sequence
CC represents a nucleic acid of a pretreated genomic DNA derived from the
CC above mentioned genes.

XX Sequence 7110 BP; 1818 A; 0 C; 1721 G; 3571 T; 0 U; 0 Other;

Query Match 54.0%; Score 21.6; DB 10; Length 7110;
Best Local Similarity 85.7%; Pred. No. 1.5e+02;

Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 ATACTTCTCTCCCAACATCATCTTAC 31
 DB 3083 AACCTAACCCCAACATCATCTTAC 3056

RESULT 14

ADBS9580/C
 ID ADS89580 standard; DNA; 7110 BP.

AC ADS89580;

DT 18-NOV-2004 (first entry)

DE Oligonucleotide of the invention SEQ ID NO:596.

KM 8a; cell proliferative disorder; breast; methylation; cytostatic;
 KM gene therapy; single nucleotide polymorphism; SNP.

OS Unidentified.

PN WO2004035803-A2.

PD 29-APR-2004.

PF 01-OCT-2003; 2003WO-EP010881.

PR 01-OCT-2002; 2002DE-01045779.

PR 07-JAN-2003; 2003DE-01000096.

PR 17-APR-2003; 2003DE-01017955.

XX (EPIC-) EPISBENOMICS AG.

PI Foekens J, Harbeck N, Koenig T, Maier S, Martens J, Model F,
 PI Nimmrich I, Rujan T, Schmitt A, Schmitt M, Look MP, Marx A;

XX WPI; 2004-348466/32.

PT Predicting responsiveness of a subject with breast cell proliferative
 PT disorder; useful for treating or differentiating breast cell
 PT proliferative disorders comprises analyzing methylation pattern of a
 PT genomic DNA from the subject.

PS Disclosure; SEQ ID NO 596; 104pp; English.

XX The invention relates to a novel method for predicting the responsiveness
 CC of a subject with a cell proliferative disorder of the breast tissues to
 CC a therapy comprising analyzing the methylation pattern of a target
 CC nucleic acid by contacting at least one of the target nucleic acids in a
 CC biological sample obtained from the subject prior to or during treatment.
 CC The method of the invention has cytostatic activity, and may have a use
 CC in gene therapy. The set of oligonucleotides comprising at least two of
 CC the oligomers are useful for detecting the cytosine methylation state
 CC and/or single nucleotide polymorphisms (SNPs) within the sequences. The
 CC methods, nucleic acid, oligonucleotide, and kit are useful for the
 CC treatment, characterization, classification and/or differentiation, of
 CC breast cell proliferative disorders. The method is also useful for
 CC predicting the responsiveness of a subject with a cell proliferative
 CC disorder of the breast tissues to a therapy. The present sequence is used
 CC in the exemplification of the invention.

XX Sequence 7110 BP; 1818 A; 0 C; 1721 G; 3571 T; 0 U; 0 Other;

Query Match 54.0%; Score 21.6; DB 13; Length 7110;
 Best Local Similarity 85.7%; Pred. No. 1.5e+02;
 Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 ATACTTCTCTCCCAACATCATCTTAC 31
 DB 3083 AACCTAACCCCAACATCATCTTAC 3056

RESULT 15

AB213750/C
 ID AB213750 standard; DNA; 1341 BP.

AC AB213750;

DT 21-JAN-2003 (first entry)

DE Arabidopsis thaliana stress regulated gene SEQ ID NO 1555.

XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.

OS Arabidopsis thaliana.

PN WO200216655-A2.

PD 28-FEB-2002.

PF 24-AUG-2001; 2001WO-US026685.

PR 24-AUG-2000; 2000US-0227866P.

PR 26-JAN-2001; 2001US-0264647P.

PR 22-JUN-2001; 2001US-0300111P.

XX (SCRI) SCRIPPS RES INST.

PA (SYGN) SYNGENTA PARTICIPATIONS AG.

PI Harper JF, Kreps J, Wang X, Zhu T;

XX WPI; 2002-304127/34.

PT Identifying a stress condition to which a plant cell has been exposed and
 PT producing plants with increased tolerance to these abiotic stresses.

XX Claim 144; SEQ ID NO 1555; 577pp + Sequence Listing; English.

XX The invention relates to identifying a stress condition to which a plant
 CC cell has been exposed, comprising: (a) contacting nucleic acid
 CC representative of expressed polynucleotides in the plant cell with an
 CC array or probes representative of the plant cell genome; and (b)
 CC detecting a profile of expressed polynucleotides in the plant cell
 CC characteristic of a stress response. The method is useful in the
 CC production of transgenic plants, cells and seeds and in producing plants
 CC with increased tolerance to abiotic stress. The present sequence is that
 CC of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used
 CC in methods of the invention. Note: The sequence data for this patent is
 CC not represented in the printed specification but is based on sequence
 CC information supplied to Derwent by the European Patent Office

XX Sequence 1341 BP; 460 A; 231 C; 319 G; 331 T; 0 U; 0 Other;

Query Match 53.5%; Score 21.4; DB 6; Length 1341;
 Best Local Similarity 71.8%; Pred. No. 1.5e+02;
 Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 AGGATCTACTCTCCCAACATCATCTTACATCA 39
 DB 288 AGTCACCTATTCACCAACATCTGTGCTCACACAA 250

Search completed: September 15, 2005, 21:33:17
 Job time : 22.3656 secs

This Page Blank (up to)

This Page Blank (up to)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 14, 2005, 19:14:19 ; Search time 6.05048 Seconds
(without alignments)
10817.505 Million cell updates/sec

Title: US-10-713-381-1_COPY_1239_1278

Perfect score: 40

Sequence: 1 aggaatccatccatcccaaacatccatccatccatcgcaac 40

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	1394	3 US-08-880-499-1	Sequence 1, Appl
2	40	100.0	1394	3 US-08-880-499-2	Sequence 2, Appl
3	20.6	51.5	1071	4 US-09-248-796A-5546	Sequence 5546, Ap
4	20.6	51.5	42075	4 US-09-949-016-14985	Sequence 14985, A
5	20.4	51.0	601	4 US-09-949-016-65770	Sequence 65770, A
6	20.4	51.0	973	4 US-09-774-528-184	Sequence 184, App
7	20.4	51.0	4201	3 US-08-945-056-4	Sequence 4, Appl
8	20.4	51.0	51770	4 US-09-949-016-13668	Sequence 13668, A
9	20.4	51.0	84571	4 US-09-949-016-17420	Sequence 17420, A
10	20.4	51.0	126200	4 US-09-949-016-11824	Sequence 11824, A
11	20.4	51.0	126200	4 US-09-949-016-13193	Sequence 13193, A
12	20.2	50.5	5099	3 US-09-610-040-5	Sequence 5, Appl
13	20.2	50.5	5099	4 US-10-267-763-5	Sequence 5, Appl
14	20	50.0	601	4 US-09-949-016-88655	Sequence 88655, A
15	20	50.0	2615	4 US-09-620-112D-22	Sequence 22, Appl
16	20	50.0	77586	4 US-09-949-016-13220	Sequence 13220, A
17	20	50.0	77586	4 US-09-949-016-13221	Sequence 13221, A
18	19.8	49.5	107	4 US-09-513-996-13202	Sequence 13202, A
19	19.8	49.5	601	4 US-09-949-016-202698	Sequence 202698, A
20	19.8	49.5	1017	4 US-09-710-279-2867	Sequence 2867, Ap
21	19.8	49.5	1164	3 US-09-134-001C-2534	Sequence 2534, Ap
22	19.8	49.5	2046	4 US-09-248-796A-6906	Sequence 6906, Ap
23	19.8	49.5	2971	4 US-09-710-279-3449	Sequence 3449, Ap
24	19.8	49.5	9366	4 US-09-949-016-17437	Sequence 17437, A
25	19.8	49.5	74096	4 US-09-949-016-11785	Sequence 11785, A
26	19.8	49.5	74097	4 US-09-949-016-16239	Sequence 16239, A
27	19.8	49.5	265038	4 US-09-949-016-15779	Sequence 15779, A

28	19.6	49.0	189	4 US-09-540-236-1562	Sequence 1562, Ap
29	19.6	49.0	601	4 US-09-949-016-83110	Sequence 83110, A
30	19.6	49.0	601	4 US-09-949-016-83111	Sequence 83111, A
31	19.6	49.0	4322	2 US-08-537-342-1	Sequence 1, Appl
32	19.6	49.0	6330	4 US-09-949-016-15506	Sequence 15506, A
33	19.6	49.0	66986	4 US-09-596-002-29	Sequence 29, Appl
34	19.6	49.0	75480	4 US-09-949-016-16090	Sequence 16090, A
35	19.6	49.0	80858	4 US-09-949-016-12659	Sequence 12659, A
36	19.6	49.0	80859	4 US-09-949-016-15715	Sequence 15715, A
37	19.6	49.0	89892	4 US-09-949-016-13667	Sequence 13667, A
38	19.6	49.0	107937	4 US-09-949-016-17192	Sequence 17192, A
39	19.6	49.0	140315	4 US-09-949-016-14141	Sequence 14141, A
40	19.4	48.5	601	4 US-09-949-016-26227	Sequence 26227, A
41	19.4	48.5	601	4 US-09-949-016-74129	Sequence 74129, A
42	19.4	48.5	601	4 US-09-949-016-74160	Sequence 74160, A
43	19.4	48.5	601	4 US-09-949-016-202761	Sequence 202761, A
44	19.4	48.5	2810	4 US-09-949-016-5447	Sequence 5447, Ap
45	19.4	48.5	2816	3 US-09-643-597-333	Sequence 333, App

ALIGNMENTS

```
RESULT 1
; US-08-880-499-1
; Sequence 1, Application US/08880499
; Patent No. 6037523
; GENERAL INFORMATION:
; APPLICANT: Albertson, Marc C.
; APPLICANT: Fox, Tim W.
; APPLICANT: Carl, Garraat W.
; APPLICANT: Huffman, Gary A.
; APPLICANT: Kendall, Timmy L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
; STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.
; STREET: Box 1000
; CITY: Johnston
; STATE: Iowa
; COUNTRY: USA
; ZIP: 50131
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/880,499
; FILING DATE: CONCURRENTLY HERewith
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Sweeney, Patricia A.
; REGISTRATION NUMBER: 32,733
; REFERENCE/DOCKET NUMBER: 0578
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (515) 248-4800
; TELEFAX: (515) 248-4844
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1394 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-880-499-1
; Query Match 100.0%; Score 40; DB 3; Length 1394;
; Best Local Similarity 100.0%; Pred. No. 4.6e-07;
; Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 AGGATACCTACTCCCAACATCATCTTACTCATGCAAC 40
|||||
Db 1239 AGGATACCTACTCCCAACATCATCTTACTCATGCAAC 1278

RESULT 2

US-08-880-499-2
; Sequence 2, Application US/08880499
; Patent No. 6037523

GENERAL INFORMATION:

APPLICANT: Albertson, Marc C.
APPLICANT: Fox, Tim W.
APPLICANT: Carl, Garraet W.
APPLICANT: Huffman, Gary A.
APPLICANT: Kendall, Timmy L.
TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION
TITLE OF INVENTION: AND METHOD OF USING SAME
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.
CITY: Johnston
STATE: Iowa
COUNTRY: USA
ZIP: 50131

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/880,499
FILING DATE: CONCURRENTLY HEREWITH
CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Sweeney, Patricia A.
REGISTRATION NUMBER: 32,733
REFERENCE/DOCKET NUMBER: 0578
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 248-4800
TELEFAX: (515) 248-4844
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1394 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

US-08-880-499-2

Query Match 100.0%; Score 40; DB 3; Length 1394;

Best Local Similarity 100.0%; Pred. No. 4.6e-07;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAACATCATCTTACTCATGCAAC 40
|||||
Db 1239 AGGATACCTACTCCCAACATCATCTTACTCATGCAAC 1278

RESULT 3

US-09-248-796A-5546
; Sequence 5546, Application US/09248796A
; Patent No. 6747137

GENERAL INFORMATION:

APPLICANT: Keith Weinstein et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
PRIOR FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13

;; PRIOR APPLICATION NUMBER: US 60/096,409
;; PRIOR FILING DATE: 1998-08-13
;; NUMBER OF SEQ ID NOS: 28208
;; SEQ ID NO 5546
;; LENGTH: 1071
;; TYPE: DNA
;; ORGANISM: Candida albicans
US-09-248-796A-5546

Query Match 51.5%; Score 20.6; DB 4; Length 1071;

Best Local Similarity 74.3%; Pred. No. 51;
Matches 26; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 4 ATACCTACTCCCAACATCATCTTACTCATGCA 38
|||||
Db 224 ATCCCACTCTCAACATCATCTTACTCATGCA 238

RESULT 4

US-09-949-016-14995/c
; Sequence 14995, Application US/09949016
; Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14995
LENGTH: 42075
TYPE: DNA
ORGANISM: Human
US-09-949-016-14995

Query Match 51.5%; Score 20.6; DB 4; Length 42075;

Best Local Similarity 85.2%; Pred. No. 1.2e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 ATACCTACTCCCAACATCATCTTACTCATGCA 30
|||||
Db 31591 ATTCTACCCCAACATCATCTTACTCATGCA 31565

RESULT 5

US-09-949-016-65770/c
; Sequence 65770, Application US/09949016
; Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 65770
LENGTH: 601

TYPE: DNA
ORGANISM: Human
US-09-949-016-65770

Query Match
Best Local Similarity 71.1%; Score 20.4; DB 4; Length 601;
Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 GGATACCTACTCCCAACATCCATCTTACTCATGCA 39
Db 414 GGATCCTCGGCGTCACGACATCCTCTGCGCTCAGGCA 377

RESULT 6
US-09-774-528-184
Sequence 184, Application US/09774528
Patent No. 6743619
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Zhou, Ping
APPLICANT: Goodrich, Ryle
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Ren, Feiyen
APPLICANT: Zhang, Jie
APPLICANT: Zhao, Qing A.
APPLICANT: Yang, Yonghong
APPLICANT: Xue, Aildong J.
APPLICANT: Wehrman, Tom
APPLICANT: Wang, Jian-Rui
APPLICANT: Wang, Dunrul
APPLICANT: Dermanac, Radoje T.
TITLE OF INVENTION: No. 6743619e1 Nucleic Acids and
FILE REFERENCE: 802
TITLE OF INVENTION: Polypeptides
CURRENT APPLICATION NUMBER: US/09/774,528
CURRENT FILING DATE: 2001-01-30
NUMBER OF SEQ ID NOS: 441
SOFTWARE: PL_FL_genes Version 2.0
SEQ ID NO 184
LENGTH: 973
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (231)..(749)
US-09-774-528-184

Query Match
Best Local Similarity 51.0%; Score 20.4; DB 4; Length 973;
Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 GGATACCTACTCCCAACATCCATCTTACTCATGCA 39
Db 626 GGAATCACTCCAAAGAAACCTCGAATCATGCA 663

RESULT 7
US-08-945-056-4
Sequence 4, Application US/08945056
Patent No. 6077994
GENERAL INFORMATION:
APPLICANT: Coupland, George M.
APPLICANT: Puteerill, Joanna J.
TITLE OF INVENTION: Genetic control of flowering
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSES: Nixon & Vanderhye PC
STREET: 8th Floor, 1100 No. 6077994th Glebe Road
CITY: Arlington
STATE: Virginia
COUNTRY: United States of America
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (RPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/945,056
FILING DATE: 20-OCT-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB95/02561
FILING DATE: 01-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9422083.7
FILING DATE: 02-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ms Mary J Wilson
REGISTRATION NUMBER: 32,955
REFERENCE/DOCKET NUMBER: 620-17
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 4201 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Arabidopsis thaliana
STRAIN: landberg erecta
POSITION IN GENOME:
MAP POSITION: chromosome 5
US-08-945-056-4

Query Match
Best Local Similarity 51.0%; Score 20.4; DB 3; Length 4201;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 GATACCTACTCCCAACATCCATCTTACT 32
Db 2631 GATACCGCTCCACACATCAACTTACT 2660

RESULT 8
US-09-949-016-13668/C
Sequence 13668, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13668
LENGTH: 51770
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(51770)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13668

Query Match
51.0%; Score 20.4; DB 4; Length 51770;

Best Local Similarity 71.1%; Pred. No. 1.5e+02;
Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 GGATACCTACTCCCAACAATCCATCTTACTCATGCAA 39
Db 9851 GGATTCCTGGCTCAAGCATCCTCGCTCAGGCAA 9814

RESULT 9
US-09-949-016-17420

; Sequence 17420, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

US-09-949-016-11824

Query Match 51.0%; Score 20.4; DB 4; Length 126200;
Best Local Similarity 71.1%; Pred. No. 1.9e+02;
Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 GGATACCTACTCCCAACAATCCATCTTACTCATGCAA 39
Db 53133 GGAAATCAACTCCAAAGGAACTTCAAAATCATGCAA 53096

RESULT 11

US-09-949-016-13193/c

; Sequence 13193, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

Query Match 51.0%; Score 20.4; DB 4; Length 126200;
Best Local Similarity 71.1%; Pred. No. 1.9e+02;
Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 GGATACCTACTCCCAACAATCCATCTTACTCATGCAA 39
Db 53133 GGAAATCAACTCCAAAGGAACTTCAAAATCATGCAA 53096

RESULT 12

US-09-610-040-5

; Sequence 5, Application US/09610040

; Patent No. 6465217

; GENERAL INFORMATION:

; APPLICANT: Boyes, Douglas

; APPLICANT: Davis, Keith

; APPLICANT: Gossamer, Jeffrey

; APPLICANT: Hamillton, Carol

; APPLICANT: Hoffman, Neil

; APPLICANT: Kloci, Andreas

; APPLICANT: Zayed, Adel

; APPLICANT: Ascenzi, Robert

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF CHORISMATE SYNTHAS

; FILE REFERENCE: 9128.14

; CURRENT FILING DATE: 2000-07-05

; NUMBER OF SEQ ID NOS: 11

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 5

; LENGTH: 5099

; TYPE: DNA

```
/ ORGANISM: Arabidopsis thaliana
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: (4045)..(4049)
/ OTHER INFORMATION: "n" indicates any nucleotide (A, C, T, or G).
US-09-610-040-5

Query Match
Best Local Similarity 50.5%; Score 20.2; DB 3; Length 5099;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY
Db 6 ACCTACTCCCAACAATCCATCTTA 30
2855 ACCTACCCCAACAATCCCTTTA 2879

RESULT 13
US-10-267-763-5
/ Sequence 5, Application US/10267763
/ Patent No. 6800459
/ GENERAL INFORMATION:
/ APPLICANT: Goriach, Jörn
/ APPLICANT: Boyes, Douglas
/ APPLICANT: Davis, Keith
/ APPLICANT: Hamilton, Carol
/ APPLICANT: Hoffman, Neil
/ APPLICANT: Kloti, Andreas
/ APPLICANT: Woessner, Jeffrey
/ APPLICANT: Zayed, Adel
/ APPLICANT: Ascenzi, Robert
/ TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
/ TITLE OF INVENTION: MODULATION OF CHORISMATE SYNTHASE AND CHORISMATE MUTASE
/ TITLE OF INVENTION: EXPRESSION OR ACTIVITY IN PLANTS
/ FILE REFERENCE: 2035DI1
/ CURRENT APPLICATION NUMBER: US/10/267,763
/ CURRENT FILING DATE: 2002-10-09
/ PRIOR APPLICATION NUMBER: 09/610,040
/ PRIOR FILING DATE: 2000-07-05
/ NUMBER OF SEQ ID NOS: 11
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 5
/ LENGTH: 5099
/ TYPE: DNA
/ ORGANISM: Arabidopsis thaliana
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: (1)...(5099)
/ OTHER INFORMATION: n = A,T,C or G
US-10-267-763-5

Query Match
Best Local Similarity 50.5%; Score 20.2; DB 4; Length 5099;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY
Db 6 ACCTACTCCCAACAATCCATCTTA 30
2855 ACCTACCCCAACAATCCCTTTA 2879

RESULT 14
US-09-949-016-88655
/ Sequence 88655, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
```

```
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 88655
/ LENGTH: 601
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-88655

Query Match
Best Local Similarity 50.0%; Score 20; DB 4; Length 601;
Matches 26; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY
Db 2 GGATACCTACTCCCAACAATCCATCTTACTGATGCA 39
293 GTATAGTCTCCTCCCAACGTTGCTTATATATAAA 330

RESULT 15
US-09-620-312D-22
/ Sequence 22, Application US/09620312D
/ Patent No. 6569662
/ GENERAL INFORMATION:
/ APPLICANT: Tang, Y. Tom
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Zhang, Jie
/ APPLICANT: Ren, Feiyan
/ APPLICANT: Chen, Rui-hong
/ APPLICANT: Zhao, Qing A.
/ APPLICANT: Wehrman, Tom
/ APPLICANT: Xue, Aigong J.
/ APPLICANT: Yang, Yonghong
/ APPLICANT: Wang, Jian-Rui
/ APPLICANT: Zhou, Ping
/ APPLICANT: Ma, Yunging
/ APPLICANT: Wang, Dunrui
/ APPLICANT: Wang, Zhiwei
/ APPLICANT: John Tillinghaest
/ APPLICANT: Drmanac, Radoje T.
/ TITLE OF INVENTION: No. 6569662el Nucleic Acids and
/ TITLE OF INVENTION: Polypeptides
/ FILE REFERENCE: 784CIP2B
/ CURRENT APPLICATION NUMBER: US/09/620,312D
/ CURRENT FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: 09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: 09/488,725
/ PRIOR FILING DATE: 2000-01-21
/ NUMBER OF SEQ ID NOS: 1105
/ SOFTWARE: pc_fl_genes Version 1.0
/ SEQ ID NO 22
/ LENGTH: 2615
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (160)..(849)
US-09-620-312D-22

Query Match
Best Local Similarity 50.0%; Score 20; DB 4; Length 2615;
Matches 26; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY
Db 4 ATACCTACTCCCAACAATCCATCTTACTGATGCA 39
1855 ATACGATTCACCAACAATCTCTCATTTGAA 1890

Search completed: September 15, 2005, 08:25:10
Job time : 9.05048 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: September 15, 2005, 07:10:11 ; Search time 26.0505 Seconds

(without alignments)
10230.248 Million cell updates/sec

Title: US-10-713-381-1_COPY_1239_1278

Perfect score: 40
Sequence: 1 aggatcaccctccccaacacatccatctcattcgcgaac 40

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7389322 seqs, 333185599 residues

Total number of hits satisfying chosen parameters: 14778644

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Labeling first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/prodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/prodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/prodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/prodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/prodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/prodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
- 14: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
- 15: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
- 16: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
- 17: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
- 18: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
- 19: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
- 20: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
- 21: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
- 22: /cgn2_6/prodata/2/pubpna/US10_NEW_PUB.seq:*
- 23: /cgn2_6/prodata/2/pubpna/US11_PUBCOMB.seq:*
- 24: /cgn2_6/prodata/2/pubpna/US11_NEW_PUB.seq:*
- 25: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq:*
- 26: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	40	20	US-10-713-381-6
2	40	100.0	50	20	US-10-713-381-5
3	40	100.0	158	20	US-10-713-381-3
4	40	100.0	255	20	US-10-713-381-9
5	40	100.0	1394	20	US-10-713-381-1
6	40	100.0	1394	20	US-10-713-381-2
7	24	60.0	85779	20	US-10-713-993-6880

8	23.3	58.0	706	18	US-10-424-599-31006	Sequence 31006, A
9	22.6	56.5	294	20	US-10-425-115-143168	Sequence 143168, A
10	22.4	315	18	US-10-424-599-83179	Sequence 83179, A	
11	22.4	56.0	226215	13	US-10-087-192-1948	Sequence 1948, App
12	21.6	54.0	4110	20	US-10-473-126-342	Sequence 342, App
13	21.6	54.0	6309	15	US-10-311-455-277	Sequence 277, App
14	21.4	53.5	1341	9	US-09-938-8428-1555	Sequence 1555, App
15	21.4	53.5	1341	11	US-09-938-8428-1555	Sequence 1555, App
16	21.4	53.5	6681	15	US-10-311-455-128	Sequence 128, App
17	21.4	53.5	6681	16	US-10-240-452-4	Sequence 4, App1
18	21.4	53.5	8168	15	US-10-311-455-901	Sequence 901, App
19	21.4	53.5	8168	18	US-10-240-452-23	Sequence 23, App1
20	21.4	53.5	3673778	16	US-10-312-841-1	Sequence 1, App1
21	21.4	53.5	3673778	16	US-10-312-841-2	Sequence 2, App1
22	21	293	18	US-10-424-599-24043	Sequence 24043, A	
23	21	52.5	931	18	US-10-424-599-20320	Sequence 20320, A
24	21	52.5	2000	17	US-10-260-238-2594	Sequence 2594, App
25	21	52.5	2792	18	US-10-041-018-361	Sequence 361, App
26	21	52.5	3117	18	US-10-041-018-362	Sequence 362, App
27	21	52.5	5507	20	US-10-473-126-198	Sequence 198, App
28	21	52.5	5507	20	US-10-473-126-344	Sequence 344, App
29	21	52.5	8404	15	US-10-311-455-1568	Sequence 1568, App
30	21	52.5	8404	18	US-10-221-714A-222	Sequence 222, App
31	21	52.5	10286	14	US-10-239-676-13	Sequence 13, App1
32	21	52.5	10286	15	US-10-240-453-21	Sequence 21, App1
33	21	52.5	63294	11	US-09-997-722-205	Sequence 205, App
34	21	52.5	106315	13	US-10-087-192-178	Sequence 178, App
35	21	52.5	2140405	13	US-10-027-632-76212	Sequence 76212, App
36	21	52.5	2140405	17	US-10-027-632-76212	Sequence 76212, App
37	20.8	52.0	2263	19	US-10-437-963-4616	Sequence 4616, App
38	20.8	52.0	3292	13	US-10-027-632-112652	Sequence 112652, App
39	20.8	52.0	3292	17	US-10-027-632-112652	Sequence 112652, App
40	20.8	52.0	6620	14	US-10-239-676-136	Sequence 136, App
41	20.8	52.0	6620	15	US-10-240-453-290	Sequence 290, App
42	20.8	52.0	513509	10	US-09-754-853A-4	Sequence 4, App1
43	20.6	51.5	970	18	US-10-424-599-91073	Sequence 91073, A
44	20.6	51.5	1103	22	US-10-499-352A-170	Sequence 170, App
45	20.6	51.5	1889	22	US-10-499-352A-178	Sequence 178, App

ALIGNMENTS

RESULT 1
US-10-713-381-6
; Sequence 6, Application US/10713381
; Publication No. US2004022131A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OR INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 6
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-6
Query Match 100.0%; Score 40; DB 20; Length 40;
Best Local Similarity 100.0%; Pred.No. 3.2e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGATCCCTACCTCCCAACATCCATCTTACATGCACAC 40

Db 1 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 40

RESULT 2

US-10-713-381-5
; Sequence 5, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-5

Query Match 100.0%; Score 40; DB 20; Length 50;
Best Local Similarity 100.0%; Pred. No. 3.3e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 40
Db 11 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 50

RESULT 3

US-10-713-381-3
; Sequence 3, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 158
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-3

Query Match 100.0%; Score 40; DB 20; Length 158;
Best Local Similarity 100.0%; Pred. No. 4.2e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 40
Db 86 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 125

RESULT 4

US-10-713-381-9
; Sequence 9, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 255
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-9

Query Match 100.0%; Score 40; DB 20; Length 255;
Best Local Similarity 100.0%; Pred. No. 4.6e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 40
Db 99 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 138

US-10-713-381-1
; Sequence 1, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-1

Query Match 100.0%; Score 40; DB 20; Length 1394;
Best Local Similarity 100.0%; Pred. No. 6.4e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 40
Db 1239 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 1278

RESULT 5

US-10-713-381-2
; Sequence 2, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.


```
/ APPLICANT: FOX, TIMOTHY W.
/ APPLICANT: GARNAT, CARL W.
/ APPLICANT: HUPFMAN, GARY
/ APPLICANT: KENDALL, TIMMY L.
/ TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
/ FILE REFERENCE: 578R
/ CURRENT APPLICATION NUMBER: US/10/713,381
/ CURRENT FILING DATE: 2003-11-14
/ PRIOR APPLICATION NUMBER: 08/880,499
/ PRIOR FILING DATE: 1997-06-23
/ NUMBER OF SEQ ID NOS: 24
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 2
/ LENGTH: 1394
/ TYPE: DNA
/ ORGANISM: Zea mays
US-10-713-381-2
```

```
Query Match          100.0%; Score 40; DB 20; Length 1394;
Best Local Similarity 100.0%; Pred. No. 6,4e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 AGGATACCTACTCCCAACAATCCATCTTACTCATGCAAC 40
DB 1239 AGGATACCTACTCCCAACAATCCATCTTACTCATGCAAC 1278
```

```
RESULT 7
US-10-719-993-6880/c
/ Sequence 6880, Application US/10719993
/ Publication No. US20040265849A1
/ GENERAL INFORMATION:
/ APPLICANT: CARGILL, Michele et al.
/ TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
/ FILE REFERENCE: CL001496
/ CURRENT APPLICATION NUMBER: US/10/719,993
/ CURRENT FILING DATE: 2003-11-24
/ NUMBER OF SEQ ID NOS: 55342
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 6880
/ LENGTH: 85779
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ NAME/KEY: misc.feature
/ LOCATION: (1)...(85779)
/ OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-719-993-6880
```

```
Query Match          60.0%; Score 24; DB 20; Length 85779;
Best Local Similarity 84.4%; Pred. No. 38;
Matches 27; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 4 ATACCTACTCCCAACAATCCATCTTACTCAT 35
DB 16082 AAACCTACTCCCAACAATCTATGTCATCAT 16051
```

```
RESULT 8
US-10-424-599-31006
/ Sequence 31006, Application US/10424599
/ Publication No. US20040031072A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J
/ APPLICANT: Kovalic, David K
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(53223)B
/ CURRENT APPLICATION NUMBER: US/10/424,599
/ NUMBER OF SEQ ID NOS: 285684
/ SEQ ID NO 83179
/ LENGTH: 315
/ TYPE: DNA
/ ORGANISM: Glycine max
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT3847_46127C.1
US-10-424-599-83179
```

```
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 285684
/ SEQ ID NO 31006
/ LENGTH: 706
/ TYPE: DNA
/ ORGANISM: Glycine max
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT3847_128002C.1
US-10-424-599-31006
```

```
Query Match          58.0%; Score 23.2; DB 18; Length 706;
Best Local Similarity 77.8%; Pred. No. 31;
Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
```

```
QY 5 TACCTACTCCCAACAATCCATCTTACTCATGCAAC 40
DB 455 TCCCTTCCCAAAACCTCCATCAACCTCATACAC 490
```

```
RESULT 9
US-10-425-115-143168
/ Sequence 143168, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(53222)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 143168
/ LENGTH: 294
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MRT4577_62049C.1
US-10-425-115-143168
```

```
Query Match          56.5%; Score 22.6; DB 20; Length 294;
Best Local Similarity 86.2%; Pred. No. 45;
Matches 25; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 4 ATACCTACTCCCAACAATCCATCTTACT 32
DB 226 ATACCTGATCCCATACATCTTACT 254
```

```
RESULT 10
US-10-424-599-83179
/ Sequence 83179, Application US/10424599
/ Publication No. US20040031072A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J
/ APPLICANT: Kovalic, David K
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(53223)B
/ CURRENT APPLICATION NUMBER: US/10/424,599
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 285684
/ SEQ ID NO 83179
/ LENGTH: 315
/ TYPE: DNA
/ ORGANISM: Glycine max
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT3847_46127C.1
US-10-424-599-83179
```

Query Match 56.0%; Score 22.4; DB 18; Length 315;
Best Local Similarity 81.2%; Pred. No. 55;
Matches 26; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 ATACCTACTCCCAACATCATCTTACTCAT 35
Db 211 AACGAACTCCTTAACATCATCTTACTCAT 242

RESULT 11

US-10-087-192-1948
; Sequence 1948, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 1948
; LENGTH: 226215
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-1948

Query Match 56.0%; Score 22.4; DB 13; Length 226215;
Best Local Similarity 72.5%; Pred. No. 2e+02;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 1 AGGATACCTACTCCCAACATCATCTTACTCATGCAC 40
Db 183206 AGAATTCCTGCTTCCAAACACAGCATCTTACTCTTTTAC 183245

RESULT 12

US-10-473-126-342/c
; Sequence 342, Application US/10473126
; Publication No. US20040234973A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Methods and nucleic acids for the analysis of hematopoietic cell
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/473,126
; PRIOR FILING DATE: 2003-09-26
; NUMBER OF SEQ ID NOS: 1258
; SEQ ID NO 342
; LENGTH: 4110
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-473-126-342

Query Match 54.0%; Score 21.6; DB 20; Length 4110;
Best Local Similarity 85.7%; Pred. No. 1.9e+02;
Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 4 ATACCTACTCCCAACATCATCTTACTCAT 31
Db 3083 AACCTTAACCCCAACATCATCTTACTCAT 3056

RESULT 13

US-10-311-455-277/c
; Sequence 277, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 277
; LENGTH: 6309
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-277

Query Match 54.0%; Score 21.6; DB 15; Length 6309;
Best Local Similarity 75.0%; Pred. No. 2.1e+02;
Matches 27; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 4 ATACCTACTCCCAACATCATCTTACTCATGCAC 39
Db 3335 ATACCTATTCCTTCCATGACCAATCTCATTAATCCAA 3300

RESULT 14

US-09-938-842A-1555/c
; Sequence 1555, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPT300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1555
; LENGTH: 1341
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1555

Query Match 53.5%; Score 21.4; DB 9; Length 1341;
Best Local Similarity 71.8%; Pred. No. 1.9e+02;
Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 1 AGGATACCTACTCCCAACATCATCTTACTCATGCA 39
Db 288 AGTCCACCTATTCCTCAACATCTGTGCTCAGACAA 250

RESULT 15

```

US-09-938-842A-1555/c
; Sequence 1555, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPT300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1555
; LENGTH: 1341
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1555

```

```

Query Match      53.5%; Score 21.4; DB 11; Length 1341;
Best Local Similarity 71.8%; Pred. No. 1.9e+02;
Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

```

```

Qy      1 AGGATACCTACTCCCAACAATCCATCTTACTCATGCAA 39
Db      288 AGTCACCTATTCTCCAAACAATCTCTGTGCTTCACACAA 250

```

```

Search completed: September 15, 2005, 20:45:52
Job time : 33.0505 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 16:07:32 ; Search time 124.754 Seconds
(without alignments)
12204.542 Million cell updates/sec

Title: US-10-713-381-1_COPY_1239_1278

Perfect score: 40

Sequence: 1 aggatactactcccaaacatccatctcactcgcgaac 40

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_ests1.*
2: gb_ests2.*
3: gb_hnc.*
4: gb_ests3.*
5: gb_ests4.*
6: gb_ests5.*
7: gb_ests6.*
8: gb_gss1.*
9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	687	9	CC656939 OGDQ20TV
2	40	100.0	915	9	CG224225 OGIAG08TV
3	40	100.0	963	9	CC656933 OGDQ20TV
4	23.8	59.5	528	7	H24137 ym50g08.r1
5	23.6	59.0	375	8	B2706605 SM41681-G
6	23.2	58.0	738	7	COL17589 GR_EB01H
7	23.2	58.0	783	9	CR181951 Reverse s
8	23.2	58.0	1018	9	CL109241 ISB1-SID1
9	23.2	58.0	1253	6	CD495749 CDA18-D05
10	23	57.5	405	8	A2883408 RPCI-23-1
11	23	57.5	798	4	BG506830 601861443
12	23	57.5	968	4	BP143726 601789606
13	22.8	57.0	771	4	BM406466 EST580781
14	22.6	56.5	410	2	AM251178 EST305315
15	22.6	56.5	560	2	BE205271 EST397947
16	22.6	56.5	624	8	CR843833 GR0AA78D
17	22.6	56.5	624	8	B2173201 CH230-369
18	22.6	56.5	644	9	CR843905 GR0AA79A
19	22.6	56.5	664	8	AZ400686 IM0167P01
20	22.6	56.5	691	2	AM687128 NF006C08R
21	22.6	56.5	736	9	CR818130 GR0AAA43D
22	22.6	56.0	1080	9	CL058859 CH216-88P
23	22.4	56.0	298	2	BB264116 BB264116
24	22.4	56.0	940	7	CK864514 AGENCOURT

c	25	22.4	56.0	940	7	CK871857	CK871857
c	26	22.4	56.0	951	7	CN317130	AGENCOURT
	27	22.2	55.5	273	1	A1206216	q227E06.x
	28	22.2	55.5	346	1	A1698201	wag7911.x
	29	22.2	55.5	398	1	AA725064	AA725064 a106n07.s
	30	22.2	55.5	417	1	A1467879	c178e09.x
	31	22.2	55.5	479	2	BF601210	266128 MA
	32	22.2	55.5	491	1	A1382034	te33g05.x
c	33	22.2	55.5	609	4	BM027586	GTT000109
	34	22.2	55.5	727	7	CK949745	CK949745 4074916 B
c	35	22.2	55.5	802	9	CG811817	FSAM35TR
c	36	22.2	55.5	940	9	CG383910	CG383910 OGDAP61TV
c	37	22.2	55.5	1090	9	CL084441	ISB1-3G16
c	38	22	55.0	207	7	CK913825	p3fmgcf.0
c	39	22	55.0	330	3	CNS09RG7	EX071651 Single re
c	40	22	55.0	467	3	CG982704	CH240_164
c	41	22	55.0	478	3	CNS09RGU	BX047050 Single re
c	42	22	55.0	506	6	CNS09070	BX036328 Single re
c	43	22	55.0	506	8	AZ021739	RPCI-23-3
c	44	22	55.0	548	3	CNS09071	BX036329 Single re
c	45	22	55.0	576	8	AQ321478	AQ321478 RCT11-10

ALIGNMENTS

RESULT 1	CC656939/c	687 bp	DNA	linear	GSS 19-JUN-2003
LOCUS	OGDQ20TV.ZM 0.7 1.5 KB	Zea mays	genomic clone	ZMM8A0554D15,	
DEFINITION	CC656939	genomic survey sequence.			
ACCESSION	CC656939	GI:32060231			
VERSION	CC656939.1				
KEYWORDS	GSS.				
SOURCE	Zea mays				
ORGANISM	Zea mays				
REFERENCE	Whitehead, C.A., Quackenbush, J., Van Aken, S., Uterback, T., Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek, R.W., Nurnberg, A., Robbins, D. and Lakey, N.				
AUTHORS	Consortium for Maize Genomics				
TITLE	Unpublished (2002)				
JOURNAL	Contact: Cathy Whitelaw				
COMMENT	Other GSSs: OGDQ20TM				
FEATURES	source	9712 Medical Center Drive, Rockville, MD 20850, USA			
	1..687	Tel: 301-838-5843			
	/organism="Zea mays"	Fax: 301-838-0208			
	/mol_type="genomic DNA"	Email: whitelaw@tigr.org			
	/strain="B73"	Seq primer: TF			
	/db_xref="taxon:4577"	Class: sheared ends.			
	/clone="ZMM8A0554D15"	Location/Qualifiers			
	/clone_1ib="ZM_0.7 1.5 KB"	1..687			
	/note="Vector: pBCK-j Site_1: HincII, 0.7-1.5 kb	/organism="Zea mays"			
	methylation filtered genomic DNA library"	/mol_type="genomic DNA"			
ORIGIN					
Query Match	100.0%;	Score 40;	DB 9;	Length 687;	
Best Local Similarity	100.0%;	Pred. No. 3.2e-05;			
Matches	40;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
1	AGGATACCTACTCCCAACATCCATCTTACTCATGCAC	40			
1					
337	AGGATACCTACTCCCAACATCCATCTTACTCATGCAC	298			

RESULT 2
LOCUS CG224225
DEFINITION CG224225 915 bp DNA linear GSS 22-AUG-2003
CG1AG08TV ZM 0.7-1.5 KB Zea mays genomic clone ZMMBMA0716B15,
genomic survey sequence.
ACCESSION CG224225
VERSION CG224225
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 915)
White, C.A., Quackenbush, J., Van Aken, S., Uteerback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
TITLE JOURNAL
COMMENT Other_GSSs: OG1AG08TV
Contact: Cathy White, Tel: 301-838-5843
Fax: 301-838-0208
Email: whitecaw@tigr.org
Seq primer: TR
Class: sheared ends.
Location/Qualifiers
1..915
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZMMBMA0716B15"
/note="Vector: pBCSK-, Site 1: HindIII, 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match 100.0%; Score 40; DB 9; Length 915;
Best Local Similarity 100.0%; Pred. No. 3.4e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAACATCCATCTTACTATGCAAC 40
|||||
Db 618 AGGATACCTACTCCCAACATCCATCTTACTATGCAAC 657
|||||

RESULT 3
LOCUS CC656933
DEFINITION CC656933 963 bp DNA linear GSS 19-JUN-2003
OGMD020TM ZM 0.7-1.5 KB Zea mays genomic clone ZMMBMA0554D15,
genomic survey sequence.
ACCESSION CC656933
VERSION CC656933.1
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 963)
White, C.A., Quackenbush, J., Van Aken, S., Uteerback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
TITLE JOURNAL
COMMENT Other_GSSs: OGMD020TV
Contact: Cathy White, Tel: 301-838-5843
Fax: 301-838-0208
Email: whitecaw@tigr.org
Seq primer: TR
Class: sheared ends.
Location/Qualifiers
1..963
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZMMBMA0554D15"
/note="Vector: pBCSK-, Site 1: HindIII, 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match 100.0%; Score 40; DB 9; Length 963;
Best Local Similarity 100.0%; Pred. No. 3.4e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAACATCCATCTTACTATGCAAC 40
|||||
Db 836 AGGATACCTACTCCCAACATCCATCTTACTATGCAAC 875
|||||

RESULT 4
LOCUS H24137
DEFINITION H24137 528 bp mRNA linear EST 06-JUL-1995
YMS0908.r1 Soares infant brain IN1B Homo sapiens cDNA clone
IMAGE:51939 5', mRNA sequence.
ACCESSION H24137
VERSION H24137.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euteria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 528)
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Ma, F.,
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.
The WashU-Merck EST Project
Unpublished (1995)
TITLE JOURNAL
COMMENT Contact: Wilson R.
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1161
Source: IMAGE Consortium, LNL
High quality sequence strops: 349
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 1161 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 349.
Location/Qualifiers
1..528
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:424875"
/db_xref="taxon:9606"
/clone_lib="IMAGE:51939"
/sex="female"
/dev_stage="73 days post natal"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares infant brain IN1B"
/note="Organ: whole brain; Vector: latmid B4; Site_1: Not
TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitecaw@tigr.org
Seq primer: TR
Class: sheared ends.
Location/Qualifiers
1..963
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZMMBMA0554D15"
/note="Vector: pBCSK-, Site 1: HindIII, 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match 100.0%; Score 40; DB 9; Length 963;
Best Local Similarity 100.0%; Pred. No. 3.4e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAACATCCATCTTACTATGCAAC 40
|||||
Db 836 AGGATACCTACTCCCAACATCCATCTTACTATGCAAC 875
|||||

RESULT 4
LOCUS H24137
DEFINITION H24137 528 bp mRNA linear EST 06-JUL-1995
YMS0908.r1 Soares infant brain IN1B Homo sapiens cDNA clone
IMAGE:51939 5', mRNA sequence.
ACCESSION H24137
VERSION H24137.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euteria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 528)
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Ma, F.,
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.
The WashU-Merck EST Project
Unpublished (1995)
TITLE JOURNAL
COMMENT Contact: Wilson R.
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1161
Source: IMAGE Consortium, LNL
High quality sequence strops: 349
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 1161 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 349.
Location/Qualifiers
1..528
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:424875"
/db_xref="taxon:9606"
/clone_lib="IMAGE:51939"
/sex="female"
/dev_stage="73 days post natal"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares infant brain IN1B"
/note="Organ: whole brain; Vector: latmid B4; Site_1: Not
TIGR

1, Site 2: Hind III; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' ACTGGAAGATTGCGCGCGGAGATTGTTTGTGTTT 3']; double-stranded cDNA was ligated to Hind III adaptors (Pharmacia), digested with Not I and directionally cloned into the Not I and Hind III sites of the lacZmid BA vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 59.5%; Score 23.8; DB 7; Length 528;
Best Local Similarity 80.0%; Pred. No. 97;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 ATACTTACTCCCAACATTCATCTTACTCATGCA 38
346 AAATCTACTCTCAACATTCATCTTACTCATGCA 380

RESULT 5
BZ706605/c 375 bp DNA linear GSS 18-MAY-2003
LOCUS SM416B1-G07.55.13.ab1 Spider Monkey genomic BAC library Ateles
DEFINITION Geoffroyi genomic, genomic survey sequence.
ACCESSION BZ706605
VERSION BZ706605.1 GI:30843142
KEYWORDS GSS.
SOURCE Ateles geoffroyi (black-handed spider monkey)
ORGANISM Ateles geoffroyi
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Ateleinae; Ateles.

REFERENCE 1 (bases 1 to 375)
AUTHORS Qian, Y., Jin, L. and Su, B.
TITLE BAC end sequences of spider monkey genomic library
JOURNAL Unpublished (2003)
COMMENT Contact: Qian Y., Jin L., Su B.
Center for Genome Information
University of Cincinnati
Kettering Lab, 3223 Eden Ave., Cincinnati, OH 45267-0056, USA
Tel: 1-513-558-6678
Fax: 1-513-558-4505
Email: eubnucmail.uc.edu
Seq primer: T7
Class: BAC ends.

FEATURES

Source

1..375
/organism="Ateles geoffroyi"
/mol_type="genomic DNA"
/strain="black-handed spider monkey"
/db_xref="taxon:9509"
/sex="Male"
/tissue_type="Skin"
/cell_type="Fibroblast"
/cell_line="AG05352"
/dev_stage="3 DA"
/clone_lib="Spider Monkey genomic BAC library"
/notes="Vector: PBACe3.6; Site 1: EcoRI; Genomic DNA was partially digested with EcoRI Vector, PBACe3.6, Recombinants were transformed into DH10B."

ORIGIN

Query Match 59.0%; Score 23.6; DB 8; Length 375;
Best Local Similarity 74.4%; Pred. No. 1.1e+02;
Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 2 GGATCTACTCTCCCAACATTCATCTTACTCATGCAAC 40
Db 255 GGATCTACTCTGTGAACCATATTTACTTAAAGCAAC 217

RESULT 6
COL17589

LOCUS COL17589 738 bp mRNA linear EST 16-JUN-2004
DEFINITION GR_EB01H15.r GR_Eb Gossypium raimondii cDNA clone GR_EB01H15 3', mRNA sequence.
ACCESSION COL17589
VERSION COL17589.1 GI:48816276
KEYWORDS EST.

SOURCE

Gossypium raimondii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eucosids II; Malvales; Malvaceae; Malvoideae; Gossypium.

REFERENCE

1 (bases 1 to 738)
Kim, H., Yu, Y., Kudrna, D., Hatfield, J., Stum, D., Mueller, C., and Udall, J. A., Rapp, R. A., Wendel, J. F., Rao, K., Soderlund, C., and Wang, R. A.

Global assembly of Cotton ESTs
Unpublished (2004)
Contact: Rod A. Wing
Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: http://genome.arizona.edu

FEATURES

Source

1..738
/organism="Gossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_EB01H15"
/tissue_type="floral"
/dev_stage="3 to +3 DPA"
/lab_host="DH10B"
/clone_lib="GR_Eb"
/note="Vector: pCW.SPORT-6.1; Site 1: NotI; Site 2: EcoRV; Library made by invitrogen with RNA supplied by Wendle lab. Directional cloned into NotI-EV. Clonies plated/picked by AGI. More glycerol clones held in -80."

ORIGIN

Query Match 58.0%; Score 23.2; DB 7; Length 738;
Best Local Similarity 77.8%; Pred. No. 1.8e+02;
Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 ATACTTACTCTCCCAACATTCATCTTACTCATGCA 39
Db 224 ATACACATCTCCATACACGCAAGTACTCATGCA 259

RESULT 7

CR181951 783 bp DNA linear GSS 06-JUL-2004
LOCUS Reverse strand read from insert in 5'HPR insertion targeting.
DEFINITION Chromosome engineering clone MHPN51p21, genomic survey sequence.
ACCESSION CR181951
VERSION CR181951.1 GI:49960800
KEYWORDS GSS; genome survey sequence; MICER.

SOURCE

Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 783)
Adams, D. J., Biggs, P. J., Cox, A. V., Davies, R. M., van der Meyden, L., Jonkers, J., Smith, J., Plumb, R. W., Taylor, R. G., Nishijima, I., Yu, Y., Rogers, J., and Bradley, A.

REFERENCE

Direct Submission
Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. http://www.sanger.ac.uk/MICER
Location/Qualifiers

FEATURES

1..783
/organism="Mus musculus"
/mol_type="genomic DNA"

/db_xref="taxon:10090"
/clone="MHPN361p21"
/clone_lib="MHPN"

ORIGIN

Query Match 58.0%; Score 23.2; DB 9; Length 783;
Best Local Similarity 77.8%; Pred. No. 1.8e+02;
Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 ATACCTACTCCCAACATCATCTTACTCATGCA 39
|||||
514 ATACCTACTCTTAGCCATGCAATCATCTTACTCCAGAA 549

RESULT 8

CL109241/c 1018 bp DNA linear GSS 05-JAN-2004
LOCUS ISB1-51D14_Sp6.1 ISB1 Xenopus tropicalis genomic clone ISB1-51D14,
DEFINITION genomic survey sequence.
ACCESSION CL109241
VERSION CL109241.1 GI:40602876
KEYWORDS GSS.

SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus; Silurana.

1 (bases 1 to 1018)
Kremetzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
Mardis, E., and Wilson, R.

A physical map of the xenopus tropicalis genome
Unpublished (2003)
Contact: Richard K Wilson

Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.edu

Insert Length: 75000 Std Error: 0.00
Seq primer: Sp6 ATTAGTGACACTATAG
Class: BAC ends

High quality sequence start: 72
High quality sequence stop: 524.
Location/Qualifiers

FEATURES

source

1..1018
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"

/db_xref="taxon:8364"

/clone="ISB1-51D14"

/clone_lib="ISB1"

/note="Vector: pBelOBAC11, ISB-1 Xenopus tropicalis BAC
Library Segment 1"

ORIGIN

Query Match 58.0%; Score 23.2; DB 9; Length 1018;
Best Local Similarity 77.8%; Pred. No. 1.9e+02;
Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 5 TACCTACTCCCAACATCATCTTCTCATGCAAC 40
|||||
Db 156 TACATCTCCCAATCATCTTATTAATGAATTAAC 121

RESULT 9

CD495749 1253 bp mRNA linear EST 12-JUN-2003
LOCUS CD418-D05_Y1d-8 SHGC-CDA Gasterosteus aculeatus cDNA clone
DEFINITION CD418-D05_3', mRNA sequence.
ACCESSION CD495749
VERSION CD495749.1 GI:31422780
KEYWORDS EST.

SOURCE Gasterosteus aculeatus (three spined stickleback)
ORGANISM Gasterosteus aculeatus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes;
Gasterosteidae; Gasterosteus.
1 (bases 1 to 1253)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Expressed sequence tags from Gasterosteus aculeatus
Unpublished (2003)
Contact: Kingsley, DM
HMI and Department of Developmental Biology
Stanford University School of Medicine
Beckman Center B300, 279 Campus Drive, Stanford, CA 94305-5329, USA
Tel: 650 725 5954
Fax: 650 725 7739
Email: kingsley@cgm.stanford.edu

Plate: 18
High quality sequence start: 15
High quality sequence stop: 667.
Location/Qualifiers

FEATURES

source

1..1253

/organism="Gasterosteus aculeatus"

/mol_type="mRNA"

/strain="Salinas river, CA"

/db_xref="taxon:69293"

/clone="CD418-D05"

/sex="mixed male and female"

/tissue_type="heads and internal organs combined"

/dev_stage="adult"

/clone_lib="SHGC-CDA"

/note="Vector: Lambda ZAP Express/PBK-CMV, Site_1: EcoRI
(5' adaptor); Site_2: XhoI (3' linker primer); The mixed
organ cDNA library was generated using the ZAP-cDNA method
by Stratagene. First strand cDNA synthesis was primed with
a 50 bp linker primer containing an oligo dt sequence
preceded by a synthetic XhoI site. 5 prime adaptors were
used containing an EcoRI cohesive end. The finished cDNAs
were inserted in to the ZAP express vector
unidirectionally in the same orientation with respect to
the lacZ promoter of PBK-CMV. An amplified library was
prepared from approximately 3 million primary clones in
the lambda ZAP Express vector. In vivo excision was then
used to generate individual PBK-CMV phagemid clones for
EST sequencing."

ORIGIN

Query Match 58.0%; Score 23.2; DB 6; Length 1253;
Best Local Similarity 77.8%; Pred. No. 2e+02;
Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 5 TACCTACTCCCAACATCATCTTACTCATGCAAC 40
|||||
Db 1047 TACCTCTTCCCAACATCATCTTACACATGCAAC 1082

RESULT 10

AZ883408 405 bp DNA linear GSS 05-MAR-2001
LOCUS RPCI-23-18912_TV RPCI-23 Mus musculus genomic clone RPCI-23-18912,
DEFINITION genomic survey sequence.
ACCESSION AZ883408
VERSION AZ883408.1 GI:13202353
KEYWORDS GSS.

SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 405)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Mouse BAC End Sequences from Library RPCI-23
Unpublished (1999)
Other GSSs: RPCI-23-18912_TV
Contact: Shaying Zhao

Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: gzhao@tigr.org
Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pdejong@mail.cno.org). Clones may be purchased from BACPAC Resources (<http://www.chori.org/bacpac/orderingframe.htm>). BAC end page: http://www.tigr.org/tcb/bac_ends/mouse/bac_end_intro.html
Plate: 189 row: L column: 2
Seq primer: SP6
Class: BAC ends.

FEATURES

Source

Location/Qualifiers
1..405

/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-189L2"
/sex="Female"
/lab_host="DH10B"
/note="Organ: Kidney/Brain; Vector: pBAC3.6; Site 1: EcoRI; Site 2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBAC3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."

ORIGIN

Query Match 57.5%; Score 23; DB 8; Length 405;
Best Local Similarity 74.4%; Pred. No. 1.9e+02;
Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAACATCATCTTACTCATGCA 39
DB 182 AGGTACCAAAATCCCAACATGCTTCTGTATCATGCA 144

RESULT 11
BG506830 798 bp mRNA linear EST 27-MAR-2001
LOCUS 601861443F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4071154 5',
DEFINITION mRNA sequence.
ACCESSION BG506830
VERSION BG506830.1 GI:13468347
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 798)
NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LMC916 row: e column: 11
High quality sequence stop: 2.
Location/Qualifiers

FEATURES

Source

1..798
/organism="Homo sapiens"
/mol_type="mRNA"

/db_xref="taxon:9606"
/clone="IMAGE:4071154"
/lab_host="DH10B (TI phage-resistant)"
/clone_lib="NIH_MGC 77"
/note="Organ: Lung; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggccatcggcc); Site 2: SfiI (ggccatcggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATTTATGCGC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGCAGCGCCGACATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.9 kb (range 0.5-4.0 kb). 12/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH-MGC Library."

ORIGIN

Query Match 57.5%; Score 23; DB 4; Length 798;
Best Local Similarity 83.9%; Pred. No. 2.2e+02;
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 10 ACTCCCAACATCATCTTACTCATGCAAC 40
DB 222 ACTCCCAACAGCCTTCTTATCATGCAATC 192

RESULT 12
BF143726 968 bp mRNA linear EST 24-OCT-2000
LOCUS 601789606F1 NCI_CGAP_Lu30 Mus musculus cDNA clone IMAGE:4020600 5',
DEFINITION mRNA sequence.
ACCESSION BF143726
VERSION BF143726.1 GI:10982766
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 968)
NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLM9274 row: k column: 01
High quality sequence start: 3
High quality sequence stop: 648.
Location/Qualifiers

FEATURES

Source

1..968
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:4020600"
/tissue_type="tumor, metastatic to mammary"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Lu30"
/note="Organ: Lung; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; transgenic model WNT-1, expression driven by MMTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies.
Investigator providing samples: Gilbert Smith, NIH"

ORIGIN

Query Match 57.5%; Score 23; DB 2; Length 968;
Best Local Similarity 74.4%; Pred. No. 2.3e+02;
Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 2 GGATACCTACTCCCAACAATCATCTTACTCATGCAAC 40
 BM406466
 DB 818 GGCTCCGCTACTCCCAACAACCAACCCCAACACCAAC 856

RESULT 13

LOCUS BM406466 771 bp mRNA linear EST 10-MAR-2003
 DEFINITION EST580781 potato roots Solanum tuberosum cDNA clone cPRO27F14 5'

ACCESSION BM406466
 VERSION BM406466.1 GI:18258084
 KEYWORDS EST.
 SOURCE Solanum tuberosum (potato)
 ORGANISM Solanum tuberosum

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanales; Solanales; Solanum.
 1 (bases 1 to 771)

AUTHORS van der Hoeven, R., Sun, H., Karamycheva, S.A., Tsai, J., Van Aken, S., Utechtback, T., Chiemingo, A., Bougri, O., Buel, C.R., Romning, C., Tankley, S. and Baker, B.

TITLE Unpublished (2001)
 JOURNAL Generation of ESTs from potato roots
 COMMENT Contact: Robin Buel

9712 Medical Center Dr, Rockville, MD 20850, USA

Email: potato-array@tigr.org
 This clone can be obtained from the University of Arizona Genomics Institute. Orders can be made through URL:

http://genome.arizona.edu/orders/
 Seq primer: T3

FEATURES

source Location/Qualifiers

1..771
 /organism="Solanum tuberosum"
 /mol_type="mRNA"
 /cultivar="Kennebec"
 /db_xref="taxon:4113"
 /clone="cPRO27F14"
 /cbase_type="roots"
 /dev_stage="in vitro grown stem cuttings"
 /lab_host="SOLR"
 /clone_lib="potato roots"
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; supplier: Cornell University, Tanksley lab.
 sequencing: The Institute for Genomic Research. Roots were isolated from in vitro grown stem cuttings on CM medium. Roots were isolated two weeks after placing the stem cuttings from in vitro grown plants on medium."

ORIGIN

Query Match 57.0%; Score 22.8; DB 4; Length 771;
 Best Local Similarity 79.4%; Pred. No. 2.6e+02;
 Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 5 TACCTACTCCCAACAATCATCTTACTCATGCA 38
 DB 649 TCCCTCCTTCACACATCCATCTCAAGCA 682

RESULT 14
 LOCUS AM257178 410 bp mRNA linear EST 20-DEC-1999
 DEFINITION EST305315 KV2 Medicago truncatula cDNA clone KV2-7D15, mRNA

ACCESSION AM257178
 VERSION AM257178.1 GI:605435
 KEYWORDS EST.
 SOURCE Medicago truncatula (barrel medic)
 ORGANISM Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

REFERENCE

1 (bases 1 to 410)
 Vandenbosch, K., Hurt, J., Moore, J., Beremand, P., Peng, H., Ellis, L., Town, C.D., Bowman, C.L., Craven, M.B., Hansen, T.S., Holt, I.E. and Fraser, C.M.

ESTs from roots of Medicago truncatula after Rhizobium inoculation
 Unpublished (1999)
 CONTACT: Vandenbosch K

Department of Plant Biology
 University of Minnesota
 220 Biosci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA
 Tel: 612 624 2755
 Fax: 612 625 1738

Email: kvandenbosch.umn.edu
 Texas A&M EST name: T115883e
 TIGR sequence name: MTAAO207K
 More information is available at: (and for clone ordering info)
 http://chryslr.tamu.edu/medicago
 Seq primer: SKmod (CPA_gAA_CTA_gtg gAT CC).

FEATURES

source Location/Qualifiers

1..410
 /organism="Medicago truncatula"
 /mol_type="mRNA"
 /cultivar="genotype A17"
 /db_xref="taxon:3880"
 /clone="KV2-7D15"
 /cbase_type="Seedling roots"
 /dev_stage="2 days post-inoculation with Sinorhizobium meliloti"
 /lab_host="E. coli strain SOLR"
 /clone_lib="KV2"
 /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into the UniZap XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-zap phage using Ex-assist helper phage and propagated in SOLR cells."

ORIGIN

Query Match 56.5%; Score 22.6; DB 2; Length 410;
 Best Local Similarity 86.2%; Pred. No. 2.8e+02;
 Matches 25; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 6 ACCTACTCCCAACAATCATCTTACTCA 34
 DB 16 ACCAATCTCCCAACAATCATCTGCTCA 44

RESULT 15

LOCUS BE205271 560 bp mRNA linear EST 05-SEP-2000
 DEFINITION EST397947 KV0 Medicago truncatula cDNA clone pKV0-21123, mRNA

ACCESSION BE205271
 VERSION BE205271.1 GI:8748567
 KEYWORDS EST.
 SOURCE Medicago truncatula (barrel medic)
 ORGANISM Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae; Medicago.

1 (bases 1 to 560)

Vandenbosch, K., Endre, G., Hur, J., Moore, J., Beremand, P., Ellis, L., Town, C.D., Bowman, C.L., Craven, M.B., Hansen, T.S., Holt, I.E. and Fraser, C.M.

ESTs from uninoculated seedling roots of Medicago truncatula
 Unpublished (1999)
 CONTACT: Vandenbosch K

Department of Plant Biology
 University of Minnesota

220 Biosci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA
 Tel: 612 624 2755
 Fax: 612 625 1738

Email: kvandend@cds.umn.edu

Texas A&M University name: T265043e

TIGR sequence name: MTGBK60TK

More information is available at:

<http://chrysis.tamu.edu/medicago>

Seq primer: Skmod (CTA gaa CTA gtc gat CC).

FEATURES

source

```
1..560
  location/Qualifiers
    /organism="Medicago truncatula"
    /mol_type="mRNA"
    /cultivar="genotype A17"
    /db_xref="taxon:3880"
    /clone="pKV0-21123"
    /tissue_type="Seedling roots"
    /dev_stage="Immediately prior to inoculation with
  Sinorhizobium meliloti (0 hour)"
    /lab_host="E.coli strain XL0LR"
    /clone_id="KV0"
    /note="Vector: Bluescript SK -; Site 1: EcoRI; Site 2:
  XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
  was directionally ligated into the Unizap XR vector from
  Stratagene and packaged using Gigapack III Gold packaging
  extracts. Plasmids containing cDNA inserts were excised
  from the recombinant lambda-Zap phage using Ex-assist
  helper phage and propagated in XL0LR cells."
```

ORIGIN

Query Match

Best Local Similarity 86.24; Score 22.6; DB 2; Length 560;

Matches 25; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY

6 ACCTACTCCCAACATCCATCTTACTCA 34

DB

24 ACCAATCCCAACATCCATCTTACTCA 52

Search completed: September 16, 2005, 08:08:40
 Job time : 131.754 secs

This Page Blank (uspto)